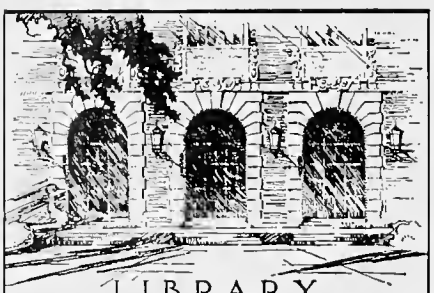

UNIVERSITY OF ILLINOIS
URBANA

AERONOMY REPORT NO. 5

AN INVESTIGATION OF IONOSPHERE-PROTONOSPHERE COUPLING

by
John E. Geisler
S. A. Bowhill

January 1, 1965

Supported by
National Science Foundation
Grant GP-866

Aeronomy Laboratory
Department of Electrical Engineering
University of Illinois
Urbana, Illinois

CITATION POLICY

Since much of the material in this report may be published in accepted aeronomic journals, it would be appreciated if persons wishing to cite work contained herein would first contact the authors to ascertain if the relevant material is part of a paper published or in process.

AERONOMY REPORT NO. 5

AN INVESTIGATION OF IONOSPHERE-PROTONOSPHERE COUPLING

by

John E. Geisler

S A Bowhill

January 1, 1965

Supported by
National Science Foundation
Grant GP-866

Aeronomy Laboratory
Department of Electrical Engineering
University of Illinois
Urbana, Illinois

ABSTRACT

The solution of a form of the steady-state continuity equation for ionized hydrogen is used in conjunction with a steady-state model of the ionospheric F2-layer to show that transfer of ionization between the F2-layer and the protonosphere at a rate sufficient to affect the structure of the mid-latitude daytime layer cannot be maintained by a diffusion process. This theory is combined with the observed diurnal variation of the electron density at the F2-layer maximum to formulate a steady-state model describing the diurnal exchange of ionization between the ionosphere and the protonosphere. The rate at which ionization diffuses downward into the ionosphere at night is shown to be dependent upon the daytime ion temperature in the topside ionosphere.

An expression is derived for the rate of heat input in the topside ionosphere which includes the contribution from photoelectrons which originate at and below the F2-layer maximum. Daytime electron and ion temperature profiles are obtained over the altitude range between 120 and 1000 km from solutions of the steady-state heat conduction equation. These are found to be in good agreement with observation.

An expression is derived for the heating of the protonosphere by photoelectrons which originate at F2-layer heights. The protonospheric temperature distribution is calculated under the assumption that cooling of the protonosphere is effected only by conduction of heat downward into the ionosphere. The effect of the downward flow of heat on the structure of the ionosphere is described for both daytime and nighttime situations.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (continued)

LIST OF ILLUSTRATIONS (continued)

LIST OF TABLES

LIST OF SYMBOLS

A = cross-sectional area of a protonospheric field tube

a_1 = inverse scale height of production rate

a_2 = inverse scale height of loss coefficient

a_3 = inverse scale height of diffusion coefficient

a_4 = inverse scale height of atomic oxygen plasma

$\overline{B}$ = magnetic field intensity

b = ratio of helium ion concentration to electron density

c = velocity of light

D = whistler dispersion (Chapter 2)

D = ambipolar diffusion coefficient (Chapter 3)

D_{jk} = binary diffusion coefficient

$\overline{E}$ = electric field intensity

E = photoelectron initial energy

e = electronic charge

$\overline{F}_j$ = force acting on a particle of gas j

F = kinetic energy flux

$F(\infty)$ = limit of F at infinite altitude

$f(E,z)$ = production rate of photoelectrons per unit interval of energy

f = frequency

f_h = local gyrofrequency

f_n = local plasma frequency

G = particle flux (electrons or atomic oxygen ions)

$G(\infty)$ = limit of G infinite altitude

LIST OF SYMBOLS (continued)

G_c = critical flux for ambipolar diffusion of atomic oxygen ions and electrons

$\bar{G}$ = particle flux (atomic hydrogen ions)

$\bar{G}_p$ = limit of $\bar{G}$ at infinite altitude

$\bar{G}_{pc}$ = critical flux for diffusion of atomic hydrogen ions

$\bar{g}$ = acceleration of gravity

$g(\alpha)$ = initial pitch angle distribution of photoelectrons

H = scale height of atomic oxygen

I = magnetic dip angle

I_ν = Bessel function of imaginary argument, order ν

J_ν = Bessel function, order ν

k = Boltzmann's constant

L = rate of cooling of electron gas

L_{ei} = rate of heat transfer from electron gas to ion gas

L_{in} = rate of heat transfer from ion gas to neutral gas

L_n = rate of heat transfer from electron gas to neutral gas

m_e = electron mass

m_+ = mean ionic mass

m_1 = mass of an oxygen atom

m_2 = mass of a helium atom

m_3 = mass of a hydrogen atom

N = total electron content of a specified volume

n = concentration of the neutral atmosphere

$n(X)$ = concentration of the constituent with chemical symbol X

LIST OF SYMBOLS (continued)

n_e = electron density

n_1 = concentration of atomic oxygen ions

n_2 = concentration of atomic helium ions

n_3 = concentration of atomic hydrogen ions

P_c = collision probability

p = pressure

p_j = partial pressure of gas j

Q = rate of heat input to electron gas

q = photoionization rate

R = photoelectron range

r = geocentric distance

r_E = radius of the earth

r_e = geocentric distance to a point in the plane of the geomagnetic equator

S = distance along a magnetic field line from the 1000 km surface to the plane of the geomagnetic equator

S_e = electron density scale factor

S_n = neutral atmosphere concentration scale factor

S_Q = heat input scale factor

s = arc length along a magnetic field line

T = temperature (when thermal equilibrium is assumed)

T_e = electron temperature

T_i = ion temperature

T_n = neutral atmosphere temperature

LIST OF SYMBOLS (continued)

t	time
T	thermal energy content of a protonospheric field tube
v	ambipolar diffusion velocity
$\lvert\bar{v}_j - \bar{v}_k\rvert$	magnitude of relative velocity of diffusion of gas j through gas k
w	instantaneous energy of a photoelectron
$\bar{w}$	mean energy of escaping photoelectrons
Z	geopotential altitude
z	altitude
z_ϵ	altitude of photoelectron escape level
α	photoelectron pitch angle
β	loss coefficient
ϵ	fraction of photoelectron initial energy given to the ambient electron gas
ρ	geomagnetic latitude
κ	charge exchange rate coefficient
λ	thermal conductivity of the electron gas
σ	scattering cross section
Φ	total flux
Ψ	energy distribution of escaping photoelectrons

1. GENERAL INTRODUCTION AND STATEMENT OF OBJECTIVES

1.1 The Neutral Atmosphere

One of the most striking features of the earth's atmosphere is the rapid diminution of density with altitude. Though the atmosphere extends for many thousands of kilometers above the earth's surface, one-half of the material of which it is composed is found below the 6 km level. At an altitude of 300 km the concentration of particles is about 10^{-15} that at ground level, corresponding to an excellent vacuum by laboratory standards.

A perpetual state of motion is the dominant characteristic of the lower atmosphere. In the extreme lower atmosphere both horizontal and vertical motions are made the more evident by the condensation of the trace constituent water vapor. Above about 15 km the density of water vapor is so low that condensation rarely occurs. The atmosphere is still characterized by winds and vertical currents, however, and this continual stirring keeps the atmosphere well mixed. Thus, water vapor is still found as high as 80 km, and the atmosphere here can still be described as air, however rarified, in the sense that the relative abundance of the constituents is very nearly the same as that near the surface of the earth.

With further increase of altitude, the chemical composition is observed to change. This change begins near the 80 km level with the dissociation of water vapor by solar radiation, a process which is responsible for the existence of the atomic hydrogen at higher altitudes. The dissociation of molecular oxygen begins near the 100 km level, and this process is the source of the atomic oxygen found at higher altitudes. These atomic constituents, along with atomic

helium (which is thought to be supplied through radioactive processes in the earth's crust), are the ultimate source of the upper ionosphere.

The chemical composition above the 100 km level is further altered by another transition that takes place at this level in that mixing of the atmosphere here gives way to diffusive separation. This does not mean that atmospheric winds cease above this level. The transition is a consequence rather of the decrease of atmospheric density: the time constant for diffusive separation of the various constituents has become shorter than the time scale of the "overturning" by vertical currents that would tend to keep the constituents mixed. The atmosphere above 120 km is to a very good approximation in a state of diffusive equilibrium. Thus each constituent is distributed independently of the others in such a way that the gradient of partial pressure counteracts the force exerted by gravity on unit volume of that constituent. The principal vertical motion in this medium characterized by several constituents in diffusive equilibrium is a diurnal one due to solar heating. This motion is negligible at the 120 km surface and increases in amplitude with altitude above this level. The change of composition at a surface of constant altitude brought about by this "diurnal breathing" of the upper atmosphere may be easily calculated, except at very high altitudes where processes other than simple diffusive separation must be considered.

A fundamental parameter of the neutral atmosphere is the kinetic temperature. Figure 1.1 shows the altitude variation of the temperature. The region above about 100 km, characterized by a temperature that increases rapidly with altitude before leveling off at a rather high value, is called the thermosphere.

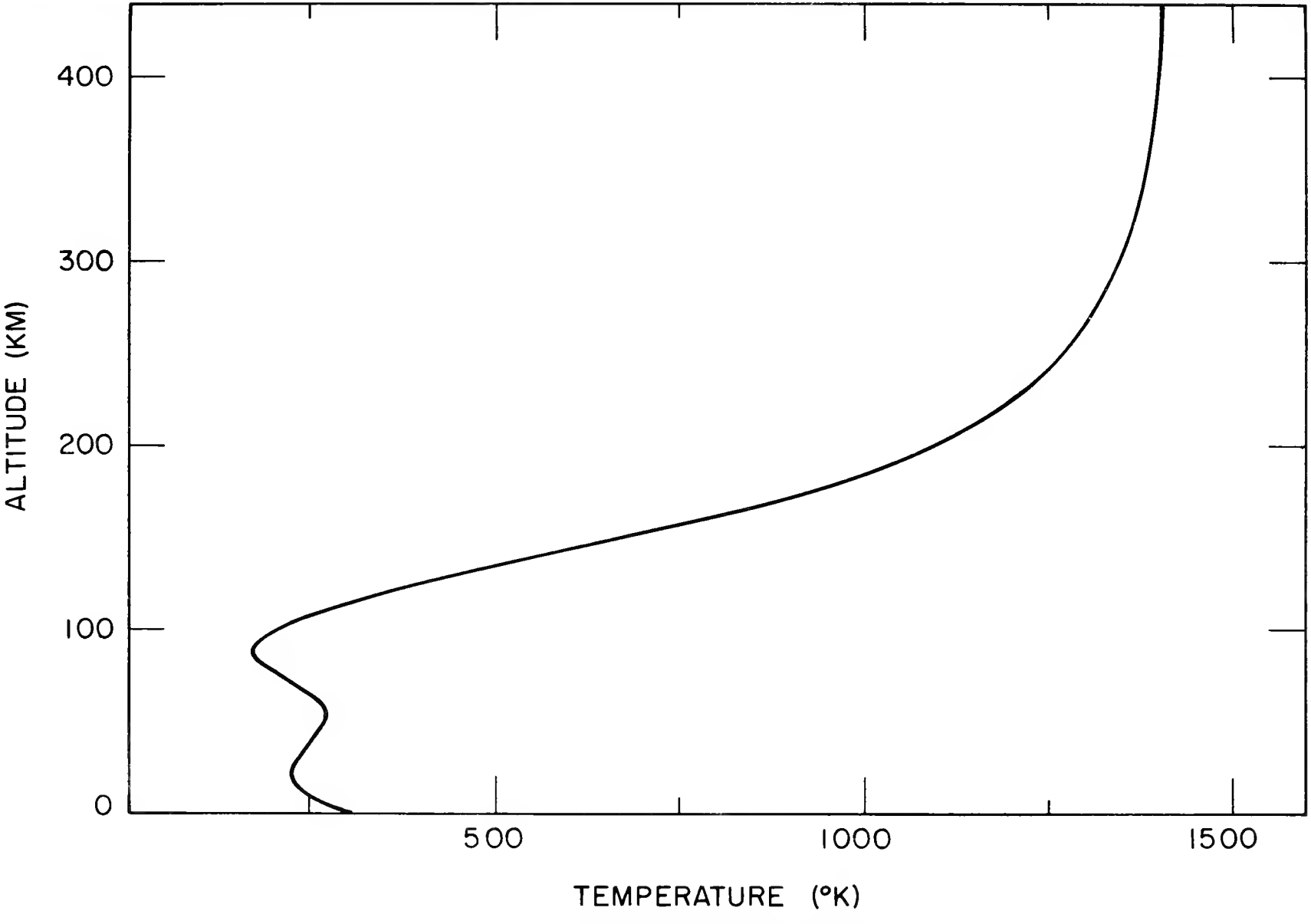

Figure 1.1 The daytime temperature of the neutral atmosphere near middle of the solar cycle. The temperature above 120 km is taken from the Harris-Priester model S = 150 for 1300 hours local time.

The thermosphere is heated by the solar ultraviolet radiation at a rate which increases rapidly with altitude above 100 km, reaching a maximum near 180 km, and decreasing rather slowly with altitude above this level. Cooling of the thermosphere is effected by infrared radiation from atomic oxygen. The cooling rate attains a maximum at the level where the atomic oxygen is found in the greatest concentration, which is near the 100 km level. As a result, downward conduction of heat plays a major role in maintaining the thermal balance of the thermosphere, and this is responsible for the large temperature gradient.

The temperature in the isothermal region above about 250 km is a strong function of time of day and epoch of solar cycle. Representative limits of the diurnal variation are $700^{\circ}K$ and $1100^{\circ}K$ at the minimum of the solar cycle and $1000^{\circ}K$ and $2000^{\circ}K$ at the maximum of the solar cycle. It is this temperature variation that results in the diurnal rise and fall of the upper atmosphere.

The temperature limits just quoted are taken from the theoretical upper atmosphere models of Harris and Priester (1962). These models are calculated from simultaneous solution of the time-dependent heat conduction equation and the equations of diffusive equilibrium for a multi-component atmosphere with time-independent composition at the 120 km level. They provide as a function of altitude above this level, the temperature of the atmosphere and the number density of O_2, N_2, O, He, and H for each hour of the day and night. Five models have been provided to cover the entire solar cycle. The absolute value and time variation of the rate of heat input as well as the ratio of number density at the 120 km level of certain constituents have been regarded as disposable parameters; they have been adjusted to give reasonable agreement with the time variation of mass density near and above the 200 km level deduced from satellite drag data.

This investigation is concerned with the behavior of the ionized component of the atmosphere in the region above about 300 km. At no time has it been found necessary to consider explicitly the effects of the neutral atmosphere below about 120 km. It has therefore been possible to make consistent use of the models of Harris and Priester throughout the course of the investigation. Unless otherwise stated, the neutral gas temperature and the number density of each neutral constituent are taken from these models.

1.2 The Ionized Component of the Atmosphere

With the possible exception of a contribution from cosmic radiation to the formation of the lowest layers, the regions of ionization that are found in the upper atmosphere owe their existence, directly or indirectly, to the ionizing power of the solar radiation. Some ionization is found at all altitudes above about 55 km, the level of the transient C-layer. Originally, it was thought that this ionization existed in the form of discrete layers, separated by regions of negligible ion content. It is now known that these "layers," (distinguished by the prefix D, E and F) are little more than slight enhancements in an electron density that increases montonically with altitude up to the F2-layer maximum. This is illustrated by the typical electron density profile shown in Figure 1.2. The original nomenclature has been retained, however, and an attempt has been made to indicate current usage by the labels appearing in the figure.

Electron density profiles up to the F2-layer maximum may be obtained through the technique of ground based ionospheric sounding. This technique makes use of the fact that the ordinary circularly polarized component of an electromagnetic wave sent vertically upward into the atmosphere will be reflected at the level where the plasma frequency is equal to that of the frequency of the wave. At

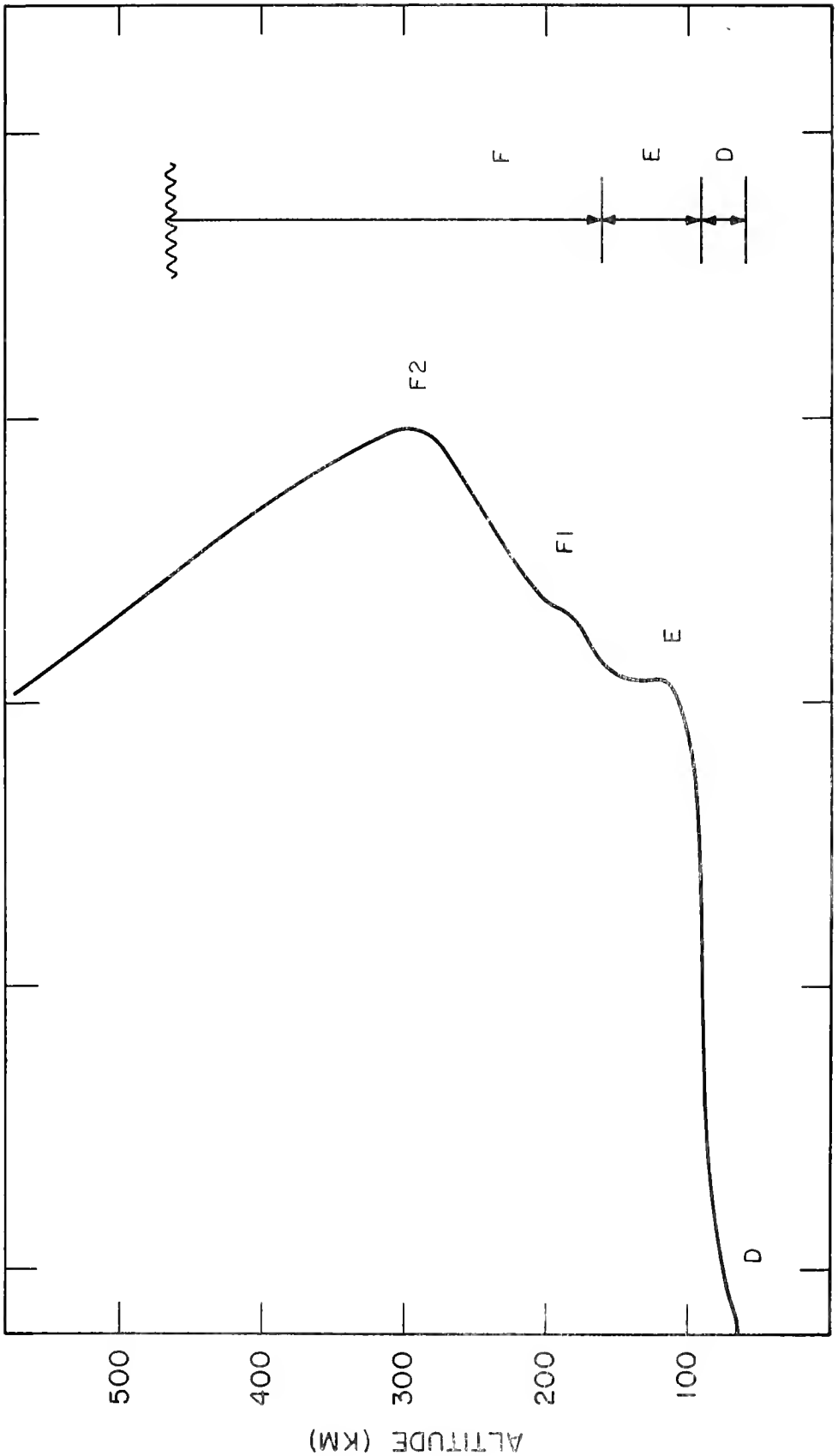

Figure 1.2 Normal daytime electron density distribution within the ionosphere near the middle of the solar cycle. The names of the layers appear next to the profile. The approximate extent of the regions is indicated at the right.

frequencies above the plasma frequency at the F2-layer maximum--the so-called critical frequency--the wave is not returned, and no information can be obtained on the electron density profile above the F2-layer maximum by this technique. This region has been given the name the topside ionosphere.

For a number of years the technique of ground-based ionospheric sounding was the only tool which could be used to obtain information about the ionosphere, and the extent and structure of the topside ionosphere remained a mystery. The breakthrough occurred in the early 1950's with the development of a theory linking the propagation of atmospheric whistlers with the electron density far above the F2-layer maximum (Storey, 1953). Modern technology has since contributed a number of other methods of studying the region, and the physics of the topside ionosphere is now probably at least as well understood as that of the lower ionosphere. More will be said of the ways and means of exploring the topside in Chapter 2.

Further discussion of the lower side of the ionosphere is beyond the scope of this investigation, except insofar as the processes occurring in this region influence the topside. In this sense the entire F2-layer must be regarded as part of the topside ionosphere, and a detailed discussion of the physics of this layer is therefore essential. This will be the subject of Chapter 3. A very brief discussion is necessary here; however, in order that the objectives of the research described in this report may be clearly set down.

The plasma of the F2-layer is one of oxygen ions and electrons. Other ionic constituents are present, but they contribute significantly to the ion density only in the lower part of the F-region. The oxygen ions are created by photoionization of atomic oxygen and disappear through chemical reactions that

will be described later. The only point to be made here is that the rate of production of these ions exceeds the rate at which they are lost throughout a region which begins at a level a short distance below the maximum and extends upward. The steady-state F2-layer is therefore characterized by a downward flow of ionization over much of its extent. This flow is one of diffusion of oxygen ions and electrons through the dominant neutral constituent present in this altitude range, atomic oxygen.

In addition to ions of atomic oxygen, the topside ionosphere is composed of helium and hydrogen ions. The chemistry of these ions will be discussed in the next chapter. All that needs be mentioned here is that, as in the case of oxygen ions, the rate of production differs from the rate of loss over much of the region which they occupy, with the result that diffusive transport is again a phenomenon to be considered.

Now throughout the topside (and the F2-layer as well) collisions between ions and neutral constituents and close ion-ion encounters are sufficiently rare that the magnetic field of the earth has a profound influence on the movement of ionization. As regards diffusion, this influence is such that diffusion across lines of magnetic force does not occur. As diffusion of ionization is characteristic of both the F2-layer and the topside ionosphere, it may be expected that the structure of this region of the ionosphere exhibits a variation with location and altitude that is strongly correlated with the form and orientation of the earth's magnetic field, and this is indeed the case.

The variation with longitude of the structure of the F2-layer and the topside is primarily a diurnal variation, and may be analyzed by consideration of the situation at a fixed longitude for different times of the day. The

latitude variation of structure is, on the other hand, not so much due to differences of solar zenith angle, as to differences of magnetic field orientation. A degree of symmetry is added if the coordinate system employed is one of geomagnetic latitude. This system is defined by the plane of symmetry which intersects at right angles the axis of the centered dipole field which most closely approximates the magnetic field of the earth. The system has been widely adopted; further details such as the transformation to geographic latitude will not be needed for this investigation. The dipole field and the geomagnetic latitude are illustrated in Figure 1.3.

Inspection of Figure 1.3 will reveal that the curvature of the lines of force is appreciable at F2-layer heights for geomagnetic latitudes below about 30°. This curvature introduces peculiarities in the structure of the F2-layer (the equatorial anomaly), and theory is correspondingly more difficult than at higher latitudes. It can also be seen from Figure 1.3 that high latitude field lines extend to very great altitudes and couple the ionosphere to the interplanetary medium. The result of this is that the ionosphere above about 65° geomagnetic latitude is frequently in a disturbed state. The existence of a belt of high energy trapped particles at this latitude is a further complication. For these reasons, the analysis will be restricted to latitudes between 30° and 65°--the so-called middle geomagnetic latitudes.

The whole of the topside ionosphere may be rather accurately described as a region of three ionic constituents (O^+, He^+, H^+) and electrons in a state of diffusive equilibrium. Unlike the situation for the neutral constituents, however, this does not imply that each constituent is distributed independently of the others. Under the influence of gravity alone, the electrons, because

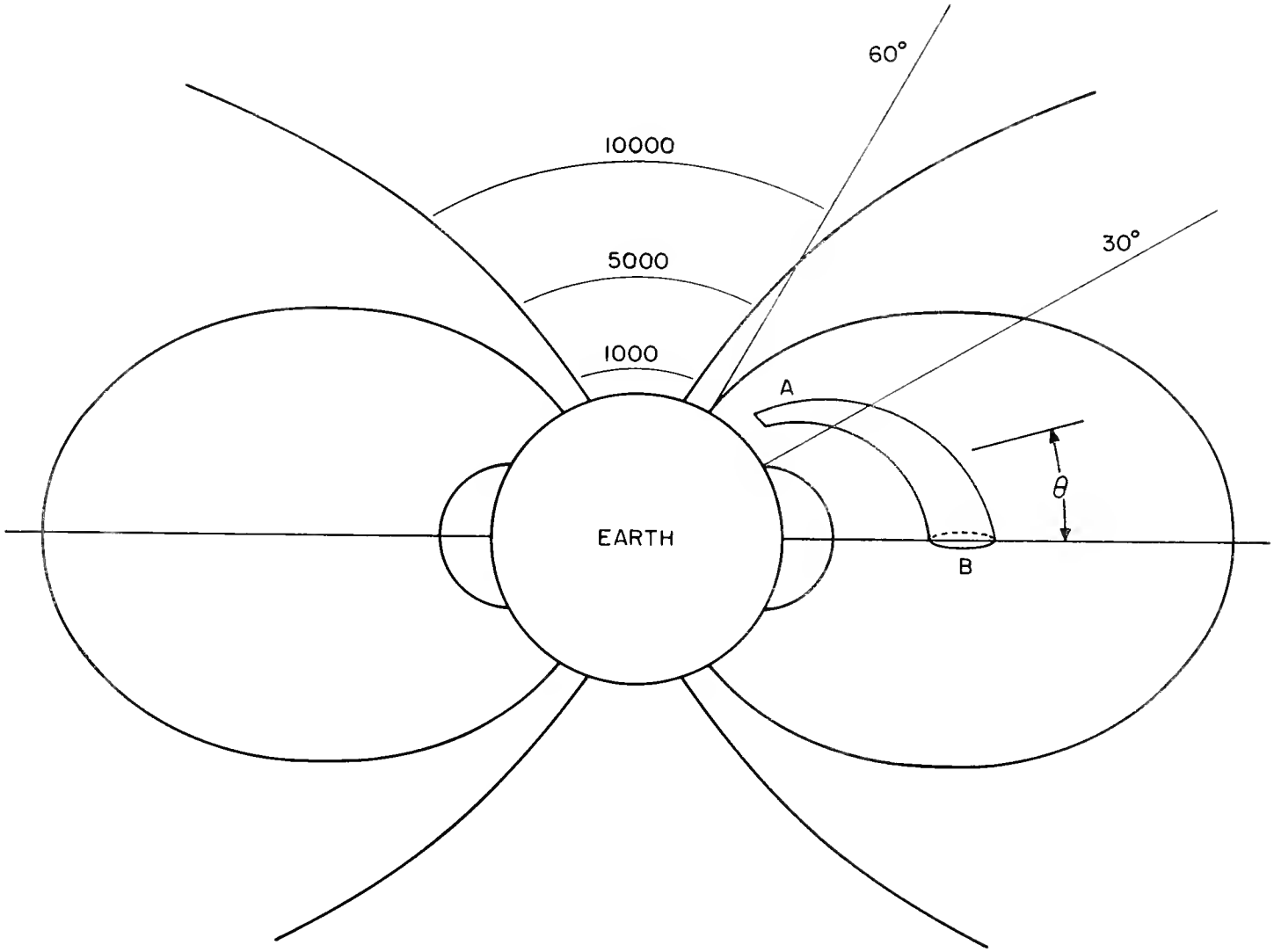

Figure 1.3 The dipole approximation to the earth's magnetic field, showing the extent of the lines of magnetic force. Altitude in kilometers is shown at the top of the figure. Geomagnetic latitude, denoted by θ, is measured from the equatorial plane. A protonospheric field tube

of their small mass, would take up a distribution decreasing with altitude much more slowly than the distribution of any of the ionic constituents. Such a situation cannot occur in the highly conducting ionospheric plasma and is prevented by the presence of a vertically directed electric field of magnitude sufficient to overcome the separative power of gravity and maintain charge neutrality. It is this electric field that couples the distribution of each ionic constituent to that of the others. The theory of diffusive equilibrium of a mixture of gases composed of charged particles will be examined in Chapter 2.

As in the case of neutral particle gases, it is characteristic of a mixture of different ion species and electrons in diffusive equilibrium that the lightest constituent eventually dominates at high altitudes. Thus, the topside ionosphere eventually becomes a region of hydrogen ions and electrons, and this region extends to very great altitudes indeed, finally merging with the interplanetary medium. The name protonosphere has been proposed for this region (Johnson, 1960). The altitude at which hydrogen first becomes the dominant ion is highly variable in time and in latitude. Generally, it occurs between about 1000 and 2000 km. The 1000 km level has been adopted as the base of the protonosphere in this investigation. It is convenient to distinguish the protonosphere as a region separate from the topside ionosphere. The nature of this distinction will be further clarified below.

The curvature of the lines of force of the earth's magnetic field must be taken into account even at middle-geomagnetic latitudes in an analysis of the protonosphere (see Figure 1.3). It was pointed out earlier that these lines of force have a constraining influence on the ionization such that diffusion of the ionization can take place only along the lines of force. Thus,

ionization can enter or leave the tube of magnetic force that extends from the base of the protonosphere to the plane of the geomagnetic equator only at the two ends (marked A and B in Figure 1.3). It is therefore essential to examine the behavior of the protonosphere with regard to behavior of the ionization contained within a tube of force. The curvature of the field lines over the range of altitude below the 1000 km level may be neglected at middle-geomagnetic latitudes, and, as the field lines are inclined rather steeply to the horizontal, it is possible to conduct the analysis of the topside ionosphere in terms of the production, loss and movement of ionization within a vertical column.

1.3 Objectives of the Investigation

Most of the assumptions fundamental to the investigation have now been stated. The behavior of the topside ionosphere may be analyzed in terms of the production, loss and movement of ionization within a vertical column extending from the lower F region to the 1000 km surface. Production and loss of ionization are negligible in the protonosphere, and a complete description of this region is obtained by specification of (1) the electron content and distribution within a tube of magnetic force bounded by the 1000 km surface and the equatorial plane and (2) the rate of change of the content due to transport of ionization through the ends of the tube. By coupling between the ionosphere and the protonosphere is meant the transport of ionization and of thermal energy across the 1000 km surface.

The object of the research described herein is to attain a quantitative understanding of the coupling between the ionosphere and the protonosphere. In particular the question to be answered is whether or not the diffusion of

ionization downward across the 1000 km surface has an appreciable effect on the morphology of the F2-layer.

1.4 Outline

There is at present no way of measuring directly the rate at which ionization is transported by diffusion processes in the upper atmosphere. However, its movements can be traced by following the variations in content of various ionospheric regions where production and loss processes are negligible. For instance, a downward flow of ionization across unit area of the 1000 km surface can only be maintained at the expense of the content of the protonospheric field tube defined by this area. Measurements of the dispersion of atmospheric whistlers have been used for some time to deduce electron densities in the plane of the geomagnetic equator, and ought therefore to provide some indication of the total electron content of a protonospheric field tube. The electron density in the equatorial plane is, however, a function of both total electron content and protonospheric temperature. The relation of the whistler dispersion to these quantities has been investigated and is described in Chapter 2.

The theory of the formation of the F2-layer is reviewed in Chapter 3, and an estimate is obtained for the magnitude of the downward flow of ionization needed to affect the structure of the layer. A theoretical treatment of ionosphere-protonosphere coupling follows in Chapter 4, where a steady-state model for coupling through exchange of ionization is set up for daytime and for nighttime conditions.

An investigation of the thermal structure of the ionosphere is described in Chapter 5. This problem has been treated previously by Hanson and Johnson (1961), by Hanson (1963), and by Dalgarno *et al.* (1963). The present treatment

differs from these in that heat transport by thermal conduction is included in the calculations. It was first noted by Hanson (1963) that a large percentage of the photoelectrons released by photoionization of atomic oxygen in the F region are able to travel vertically upward through the topside and into the protonosphere, heating these regions as they lose energy. A model for this heating process has been set up, and detailed calculations for the rate of heat input to the topside are included in this chapter.

This analysis is extended in Chapter 6 to derive the rate of heat input to a protonospheric field tube. The only mechanism for cooling the protonosphere is conduction of heat downward through the 1000 km surface. Estimates of the thermal energy content of a protonospheric field tube and the rate at which this tube cools by night are obtained. The effect of this downward flux of heat on the thermal structure of the nighttime ionosphere is described.

2. THE TOPSIDE IONOSPHERE AND THE PROTONOSPHERE

2.1 Introduction

In the opening chapter it was remarked that the topside ionosphere consists of a mixture of electrons and of ions of oxygen, helium and hydrogen which, to a first approximation, may be considered to be in a state of diffusive equilibrium. In this chapter a better approximation to the distribution of the ion species present in the topside will be formulated.

Departures from a state of diffusive equilibrium are brought about over some regions of the topside by the control of chemical processes and over other regions by diffusive transport of ionization. It is convenient to distinguish the regimes of departure from diffusive equilibrium by the following terminology:

(i) If an ionic constituent is produced by and lost through chemical processes, then the region over which the rate of loss per unit volume is equal to the rate of production per unit volume is one of chemical equilibrium. For an ionic constituent produced by photoionization and lost through a chemical process, the region over which the two rates are equal is one of photochemical equilibrium.

(ii) Diffusive transport of a particular ionic constituent may result either from inequality of production and loss rate over a region of the topside or from a source of ionization external to the topside. If the departure of the distribution of this constituent from that of diffusive equilibrium required to maintain this flow is small, then the region is one of fast diffusion. If the required distribution differs markedly from that of diffusive equilibrium, the region is one of slow diffusion.

Since any departure from a distribution of diffusive equilibrium is attended by a diffusive flow, it follows that regions of chemical and photochemical equilibrium are also regions of either slow diffusion or fast diffusion.

The behavior of a plasma composed of electrons and three species of ions is very complicated, and several assumptions are necessary to reduce the problem to manageable proportions. These assumptions will be set forth in Section 2.3, where the equations governing the ion distributions are presented. The principal assumption, which may be stated at this time, is that the explicit time dependence of all physical quantities will be ignored. Where it is necessary to consider phenomena which result from differences between the daytime and nighttime structure of the topside (as in Chapter 4), the analysis will be carried out using a steady-state model for each of these periods.

A short discussion of the production and loss processes which govern the ion species present in the topside follows in Section 2.2. The theory of diffusion processes occurring in the topside is presented in Section 2.3. With the aid of the equations derived in this section, the structure of the topside and of the protonosphere may be analyzed; this is done in Sections 2.4 and 2.5, respectively. Section 2.6 follows with a review of the techniques of measuring the composition and temperature of the topside. The chapter is concluded with a section of the relation between the dispersion of a whistler and the electron temperature in the protonosphere.

It is convenient to terminate this introductory section with a short digression on the subject of notation and terminology. The symbol n will be used to denote the number of particles per unit volume. Where these particles

are molecules, atoms or ions, the conventional chemical symbol will appear in parenthesis after this symbol, as, for example, n(O) for atomic oxygen and $n(O^+)$ for atomic oxygen ions. The term concentration will be used when referring to the number of any of these constituents per unit volume. An exception to this convention occurs in the treatment of a mixture of gases, where on occasion, subscripted notation is used to distinguish between the gases present. Where the particles are electrons, the symbol n_e will be used, and this will be called the electron density. It should be emphasized that this is a number density and not a mass density. Exceptions to this terminology for the electrons may occur, as in the phrase, "equality in the concentration of electrons and oxygen ions."

2.2 Chemistry of the Ions Present in the Ionosphere

The presence of oxygen ions is due to photoionization of atomic oxygen by solar radiation in the far ultraviolet and the X-ray region of the spectrum. The ionization is produced at a rate of q per unit volume which, at F2-layer heights and above, is very nearly proportional to the concentration of atomic oxygen. Thus, the production rate over this region exhibits an altitude dependence which is the same as that of atomic oxygen.

Loss of oxygen ions takes place through the reaction

$$O^+ + XY \longrightarrow XO^+ + Y, \qquad (2.1)$$

followed by

$$XO^+ + e \longrightarrow X + O. \qquad (2.2)$$

where XY is a molecular species (molecular oxygen or molecular nitrogen); the primes are used to denote the fact that the end products may be in an excited state. The first reaction is one of ion-atom interchange, while the second is one of dissociative recombination. Near the F2-layer and above, the first reaction proceeds at the slower rate, and the rate of loss is therefore jointly proportional to $n(O^+)$ and $n(XY)$. In this range of altitudes the concentration of molecular ions is much less than that of oxygen ions and it follows that $n(O^+) = n_e$. It is therefore possible to set the loss rate equal to βn_e, where β is the loss coefficient and has the altitude dependence of the molecular species which is dominant in the ion-atom interchange reaction.

The region of photochemical equilibrium for oxygen ions is located some distance below the F2-layer maximum. The increase with altitude of the effectiveness of diffusion over loss as the process preventing accumulation of ionization is responsible for the existence of this maximum. A short distance above the maximum there is a transition from a region of slow diffusion to one of fast diffusion. Thus, even though the topside ionosphere is characterized by the presence of a rather large downward flow of oxygen ions, the distribution of these ions is very close to one of diffusive equilibrium.

The existence of an ionized medium extending upward over a distance of several earth radii was first inferred from the propagation of atmospheric whistlers (Storey, 1953). This medium is one of hydrogen ions and electrons which are created by the chemical process of charge exchange over a region near and above the F2-layer maximum (Dungey, 1955; Johnson, 1960). As the difference between the ionization potentials of atomic oxygen and atomic

ydrogen is a small fraction of the thermal energy, the charge exchange reaction

$$H + O^{+} \rightleftharpoons H^{+} + O \qquad (2.3)$$

roceeds almost equally rapidly in either direction and thus accounts for both he production and loss of hydrogen ions. The rate of production of these ions s jointly proportional to the concentration of neutral hydrogen and atomic xygen ions with constant of proportionality K (the rate coefficient), while he rate of loss is jointly proportional to the concentrations of hydrogen ions nd atomic oxygen. The coefficient for the loss process is just $8K/9$ (Rapp, 963). It follows that the distribution of hydrogen ions in the region of hemical equilibrium is given by

$$n(H^{+}) = \frac{9}{8} \frac{n(H)}{n(O)} n(O^{+}). \qquad (2.4)$$

A fact which will be discussed in more detail later is that the distribution iven by Equation (2.4) is the same as that which is adopted by hydrogen ions n diffusive equilibrium as a minor constituent in the presence of oxygen ions. ince hydrogen ions are in fact a minor constituent of the topside in the region f chemical equilibrium, it follows that this region is one of fast diffusion. he region of chemical equilibrium extends upward to an altitude of about 700 km, here a transition to a region of slow diffusion occurs. This region continues p to the base of the protonosphere and there merges into a region of fast dif- usion. The diffusion process in the slow diffusion region is one of hydrogen ons through a mixture of oxygen ions and helium ions. The physical situation ere differs considerably from that in the slow diffusion region for oxygen ions

near the F2-layer maximum, where the diffusion process is one of oxygen ions through atomic oxygen.

The realization that ionized helium is present in the topside ionosphere in sufficient quantity that it is the dominant constituent over part of the region has come about comparatively recently. This was suggested by Nicolet (1961), and the first experimental confirmation was reported by Hanson (1962) from an analysis of data provided by a rocket-borne measurement of topside ion density (Hale, 1961). Subsequent measurements by rocket and satellite have further confirmed the existence of the helium ion layer.

The helium ions are created by photoionization of atomic helium and disappear through ion-atom interchange according to the reaction

$$He^+ + O_2 \longrightarrow HeO^+ + O', \qquad (2.5)$$

followed by radiative dissociation (Bates and Patterson, 1962). Calculation of the distribution of helium ions by these authors indicates that the region of photochemical equilibrium extends to about 350 km. The downward flow of helium ions is rather small, and the region above this altitude is essentially one of fast diffusion.

2.3 The Diffusion Equations

The diffusion of one gas through another is due to the relative thermal motion of the particles of the two gases. The process whereby gas j diffuses through gas k may be affected by external forces acting on the molecules of gas j and is opposed by collisions between the particles of gas j and those of gas k. Viewed macroscopically, the effect of collisions in opposing diffusion is analogous to frictional drag and is jointly proportional to the product of the

concentration of the two gases and their mean relative velocity. Thus, the equation of motion of the gas j in the presence of a mixture of gases may be written (Chapman and Cowling, 1958):

$$n_j m_j \frac{D_j \bar{v}_j}{Dt} = n_j \bar{F}_j - \nabla p_j + \sum_k \Theta_{jk} n_j n_k (\bar{v}_j - \bar{v}_k), \qquad (2.6)$$

where D_j/Dt is the time derivative following the motion of the gas. In this equation, n_j and m_j are the concentration and mass of a particle of gas j, which, under the action of a force $\bar{F}_j$ and a gradient of partial pressure ∇p_j, moves through the gas k with relative velocity $(\bar{v}_j - \bar{v}_k)$. The terms under the summation will be referred to as the drag terms. The coefficient Θ_{jk} is related to the diffusion coefficient D_{jk} according to

$$\Theta_{jk} = \frac{kT}{n D_{jk}} , \qquad (2.7)$$

where k is Boltzmann's constant, T is the temperature and $n = (n_j + n_k)$.

It was pointed out in the preceding section that diffusion of oxygen ions and electrons through atomic oxygen is a process fundamental to the formation of the F2-layer. Further discussion of this particular diffusion process will be deferred until the discussion of the structure of the F2-layer in Chapter 3, and attention will be directed in this section to processes of charged-particle diffusion occurring in the topside.

In the derivation to follow it is convenient to distinguish the ion concentrations by subscripts rather than by the conventional notation $n(O^+)$, $n(He^+)$ and $n(H^+)$; the symbols n_1, n_2 and n_3 will be used, respectively, to denote these concentrations. The symbol n_e will, however, be retained for

the electron density. It will be assumed that the ionic constituents have a common temperature T_i, which may differ from the electron temperature T_e.

As stated previously, the time dependence will be neglected. Under this assumption, the equation governing the distribution of ion species j is

$$\nabla p_j - n_j \bar{F}_j = \sum_k \Theta_{jk} n_j n_k (\bar{v}_j - \bar{v}_k) + \Theta_{je} n_j n_e (\bar{v}_j - \bar{v}_e), \tag{2.8}$$

with a similar equation for the electrons. The diffusion coefficient for the ion mixture is given by (Chapman and Cowling, 1958).

$$D_{jk} = \frac{3}{4n} \left\{ \frac{2}{\pi} \left(\frac{m_j m_k}{m_j + m_k} \right)^{-1} \right\}^{1/2} \frac{(kT_i)^{5/2}}{e^4 \log \Lambda}, \tag{2.9}$$

where

$$\Lambda = 1 + \left(\frac{4\sigma k T_i}{e^2} \right)^2, \tag{2.10}$$

and e is the magnitude of the electronic charge. The quantity d in Equation (2.10) is taken to be the interparticle distance by Chapman and Cowling. It should probably be taken as the Debye length (Ferraro, 1964), but the value of the diffusion coefficient is scarcely affected by this choice, since d appears only in the logarithmic factor. A further difficulty arises because the theory of Chapman and Cowling is based on the assumption that the temperature is the same for all constituents, and the way in which the ion and

electron temperatures should enter the expression for the electron-ion diffusion coefficient D_{je} is not immediately obvious. This presents no difficulty in practice, since, as will now be shown, the contribution of the electron drag term to the right-hand side of the Equation (2.8) is negligible.

Consider first the ratio of $|\bar{v}_j - \bar{v}_e|$ to $|\bar{v}_j - \bar{v}_k|$. If this ratio is large the motion is one of an ion gas composed of ion species j and k drifting through a gas of electrons--that is, a current. The steady-state currents which are present in the topside are very small, however, and it may be concluded that this ratio is close to unity. With the aid of Equation (2.9), it can then be seen that the ratio of the electron drag term to that of the ion species k is not larger than the quantity

$$\left\{ \frac{m_e}{m_k} \left(\frac{m_j + m_k}{m_j} \right) \right\}^{1/2} \frac{n_e}{n_k} .$$

For the mixture of ions considered here, the ratio m_j/m_k is at most 16, while the ratio m_e/m_k is at most 1/1843. Thus, the ratio of drag terms is no larger than $10^{-1}(n_e/n_k)$. Only if $n_k \ll n_e$ for all k would it be possible for the electron drag term to dominate. However, $n_k \ll n_e$ for all k implies that $n_j \approx n_e$, and if large currents do not exist, then it must be so in this case that $(\bar{v}_j - \bar{v}_e) \approx 0$. The result for this case is that the entire left-hand side of Equation (2.8) is negligibly small, and the distribution of ion species j is close to that of diffusive equilibrium.

Therefore, it may be concluded that the electrons offer negligible resistance to the diffusion of one ion species through another. If this is so,

however, then it must be concluded also that the diffusing ions have little or no effect on the distribution that the electrons take up. Thus, all of the drag terms in the equation for the electron distribution are negligible, and the equations governing the distribution of the ionization in the topside are

$$\nabla p_J - n_J \bar{F}_J = \sum_k \Theta_{Jk} n_j n_k (\bar{v}_J - \bar{v}_k), \tag{2.11}$$

$$\nabla p_e - n_e \bar{F}_e = 0. \tag{2.12}$$

The force terms must now be specified, and it is here that further simplifying assumptions are necessary. The topside is permeated by a magnetic field $\bar{B}$, and the gravitational force which tends to separate the ions and the electrons is opposed by an electric field $\bar{E}$. Thus, quite generally, the force $\bar{F}_J$ acting on an ion of the species j is given by

$$\bar{F}_J = e\left(\bar{E} + \frac{\bar{v}_J}{c} \times \bar{B}\right) - m_J \bar{g}, \tag{2.13}$$

where c is the velocity of light and $\bar{g}$ is the acceleration of gravity, with a similar relation for $\bar{F}_e$. A crossed electric and magnetic field gives rise to a drift of ionization in the direction $\bar{E} \times \bar{B}$ (the Hall drift). The presence of the electric field implies an imbalance in the pressure gradient and the gravitational force, and the resulting force field crossed with the magnetic field drives a current in the direction of $\bar{E} \times \bar{B}$ (the Hall current). The magnitudes of the Hall drift and the Hall current are rather small, and detailed consideration of these phenomena is not essential to an understanding of the more

important steady-state features of the topside. These features are brought about by diffusion, which as remarked earlier, takes place freely only along the direction of the magnetic field.

It will now be assumed that the topside ionosphere is plane-stratified and permeated by a uniform magnetic field whose direction makes an angle I with the horizontal. As the only motion to be analyzed here is one of diffusion along the lines of force of this field, it will be further assumed that the relative velocities in Equation (2.11) have no component perpendicular to the lines of force. The magnetic force term in Equation (2.13) therefore vanishes. Equation (2.11) must be modified to conform to this constraint imposed on the velocities. This may be accomplished by taking the scalar product of both sides of the equation with a unit vector in the direction of the magnetic field. As the force terms (electric field, gravity) and the pressure gradient are all directed along the vertical, the sole effect of this operation is to introduce a term $\sin I$ on the left-hand side of Equation (2.11). It is desirable to alter this equation further to one which expresses the relation between the forces acting on the ionization contained in unit vertical column and the component of relative velocity along that column. This may be accomplished by dividing the velocities on the right-hand side of the equation by $\sin I$.

Thus, the equations governing the distribution of ionization which is constrained to move only along lines of magnetic force in a plane-stratified model of the topside ionosphere are obtained.

$$\frac{dp_j}{dz} = n_j (eE - m_j g) - k T_j \sum_k \frac{n_j n_k (v_j - v_k)}{n_j n_k D_{jk} \sin^2 I} \qquad (2.14)$$

$$\frac{dp_e}{dz} + n_e (eE + m_e g) = 0, \qquad (2.15)$$

where z is the vertical coordinate, and the relative velocities are along the vertical direction.

2.4 The Structure of the Topside Ionosphere

Apart from the regions of chemical and photochemical equilibrium, the only portion of the topside where the distribution of a constituent differs appreciably from that of diffusive equilibrium is over a region of slow diffusion. As pointed out in Section 2.2, there are only two such regions: one for oxygen ions located very close to the F2-layer maximum, and one for hydrogen ions between 700 and the base of the protonosphere. The region of slow diffusion for hydrogen ions will be considered in detail in Chapter 4, as it is this region that regulates the diffusive coupling between the ionosphere and the protonosphere. With the exception of this region for the hydrogen ions, the structure of the topside is very accurately represented by a mixture of electrons and three ionic constituents each in diffusive equilibrium above their respective levels of chemical and photochemical equilibrium.

The distributions taken up by each of the ionic constituents in a mixture of ions and electrons in diffusive equilibrium in a gravitational field have been considered by a number of investigators, although the work of Mange (1960) is most often cited. The treatment which is given here follows most closely that of Hanson (1962). It will be assumed that the region under consideration is isothermal; however, allowance will be made for the possibility that the temperature common to the ion species present may differ from the electron temperature.

With the foregoing assumptions regarding the temperature and the absence of a diffusive flow, Equations (2.14) and (2.15) take the form

$$\frac{dn_j}{dz} = \frac{n_j}{kT_i}(eE - m_j g), \tag{2.16}$$

$$\frac{dn_e}{dz} = -\frac{n_e}{kT_e}(eE + m_e g), \tag{2.17}$$

where use has been made of the ideal gas law relating the partial pressure to the concentration of each constituent and the temperature. Equation (2.16) governs the distribution of oxygen ions, helium ions or hydrogen ions, according as $j = 1$, 2 or 3. The distribution of each of these species is dependent upon that of the others through the electric field term. This field, it was pointed out earlier, is required to maintain charge neutrality in the presence of the separative action of gravity.

It follows from the condition of charge neutrality that the sum of the left-hand sides of Equations (2.16) is equal to the left-hand side of Equation (2.17). The equation resulting from this condition provides an explicit relation for the electric field:

$$eE = \left[\sum_J \frac{m_j n_j}{n_J} + m_e\right] \frac{g}{(1+T_i/T_e)} . \tag{2.18}$$

It is convenient to define a mean ionic mass by

$$m_+ = \sum_j \left(\frac{m_j n_j}{n_j}\right) . \tag{2.19}$$

Evidently, $m_e \ll m_-$ and may therefore be neglected in Equation (2.18). Upon substitution for the electric field from Equation (2.18), the equation governing the distribution of the ions becomes

$$\frac{dn_J}{dz} = n_J \left\{ \frac{m_-}{(1 + T_i/T_e)} - m_J \right\} \frac{g}{kT_i} \quad . \tag{2.20}$$

The distribution of ion species j may be found by integration of this equation. There is no further need for the equation governing the electron distribution, since the electron concentration at any level is equal to the sum of the ion concentrations.

A complication that arises in the integration of Equation (2.20) is the variation of g with altitude over the topside. This difficulty may be avoided by transformation to a variable Z known as the geopotential altitude (sometimes also referred to as the reduced altitude)

$$Z = \left(\frac{r_E}{r_E + z} \right) z, \tag{2.21}$$

where r_E is the radius of the earth and z is altitude. This transformation has the property that

$$\frac{dZ}{dz} \quad \frac{g}{g_E} \quad . \tag{2.22}$$

Upon transformation to the geopotential altitude, Equation (2.20) may be

integrated to give

$$n_j(Z) = n_j'(Z) \exp\left[\frac{g_o}{kT_i(1 + T_i/T_e)} \int_{Z_o}^{Z} m_+(Z)\, dZ\right], \qquad (2.23)$$

where

$$n_j'(Z) = n_{jo} \exp\left[-\frac{m_j g_o}{kT_i}(Z-Z_o)\right]. \qquad (2.24)$$

The quantity n_{jo} is the concentration of ion species j that would be found at the surface of the earth if this constituent were in diffusive equilibrium down to this level. This, of course, is not the case, but n_{jo} may also be interpreted as the concentration at any other level provided that the geopotential altitude at this level is substituted for Z_o in Equations (2.23) and (2.24).

The integral in Equation (2.23) may be evaluated by a procedure due to Hanson (1962), but details will not be given here. The final result may be put into the form

$$n_j(Z) = n_j'(Z) \left\{\frac{\sum_\ell n_{\ell o}}{\sum_\ell n_\ell'(Z)}\right\}^{T_e/(T_i+T_e)} \qquad (2.25)$$

The dependence of $n_j(Z)$ on the distribution of the remaining constituents is clearly illustrated here. Given the electron and ion temperature and the concentration of each of the constituents at any level where all of them are in a state of diffusive equilibrium, the distribution of any one of these constituents over the region where diffusive equilibrium prevails for that constituent may be found from Equation (2.25).

It will be shown in Chapter 5 that there is a considerable difference between electron and ion temperature over much of the topside. By a procedure described in that chapter, the calculated ion and electron temperatures are incorporated into the model of the topside adopted for this calculation. The altitude variation of the ionic constituents of this model is given by Equation (2.25), and it is for this reason that T_e and T_i have been retained as separate entities in the derivation of this equation. In the remainder of this chapter and in Chapter 4, it will be assumed that ion and electron temperatures are equal over the topside. The effect of this assumption on the results of Chapter 4 will be discussed at the end of that chapter. As regards the analysis of the relation between the dispersion of a whistler and the electron temperature in the protonosphere, to appear in Section 2.7, this assumption has no effect, as ion and electron temperatures are substantially equal in the protonosphere.

The distribution of ionic constituents and electrons through the topside that obtains when $T_e = T_i = 1250^{\circ}K$ and the concentration of each constituent at the 500 km level is as given in Table 2.1 is illustrated in Figure 2.1 by the continuous lines. The ordinate is true altitude; the non-linear scale is due to the fact that the distributions are calculated and plotted in terms of geopotential altitude.

Ion	j	n
O^+	1	2×10^5
He^+	2	4×10^3
H^+	3	1.5×10^2

Table 2.1. Ion concentration at 500 km for model shown in Figure 2.1.

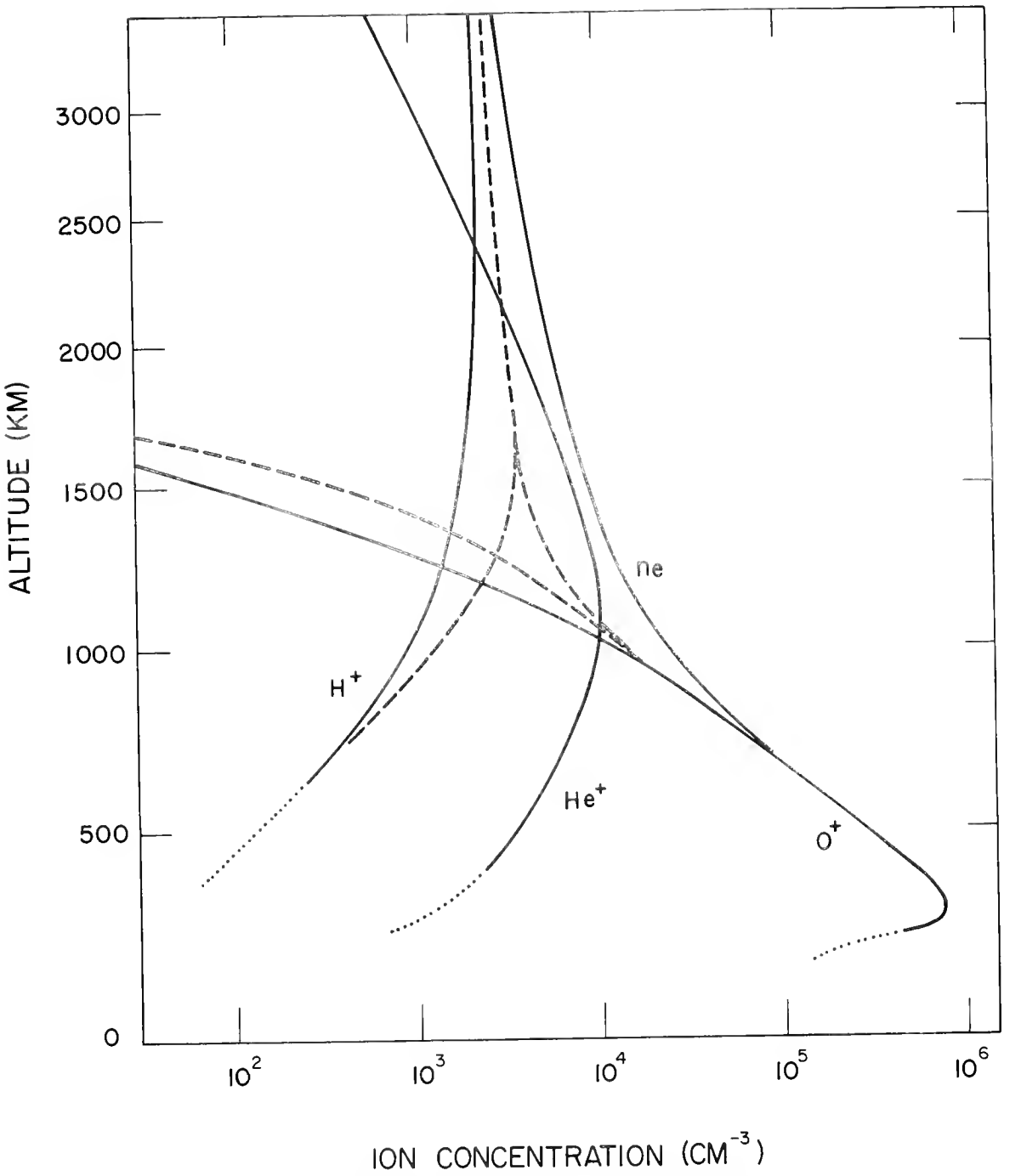

Figure 2.1 Diffusive equilibrium model of the topside ionosphere when T = 1250°K and the ion concentrations at 500 km are assumed to be those listed in Table 2.1. The broken lines show the distribution of ionization in the absence of helium ions. The distributions have been extrapolated below the 500 km level; dotted lines indicate approximate distributions over regions of chemical and photochemical equilibrium.

The distributions seen here differ markedly from those which would be taken up by a mixture of three neutral constituents in diffusive equilibrium in that the concentration of each neutral constituent would decrease monotonically with altitude. This difference is brought about by the electric field in the manner described below.

Consider first the region over which both helium ions and hydrogen ions are minor constituents. The electric field will be such as to maintain approximate equality in the concentration of electrons and oxygen ions, and the concentration of oxygen ions will decrease more slowly with altitude than it would under the influence of gravity alone. The magnitude of the upward force which the field exerts on an oxygen ion is, from Equation (2.18)

$$eE \approx m_1 g/2,$$

that is, about one half that of the gravitational force. This same field acts on the minor constituents present; for ions of helium and hydrogen, the force due to the field exceeds that of gravitation by factors of 2 and 8, respectively. The net force on these ions is therefore in the upward direction, and this is the reason why the concentration of these constituents must increase with altitude.

The situation just described is reversed at higher altitudes, where oxygen ions become the minor constituent. The electric field required to maintain charge neutrality among the now dominant light ions and the electrons is weaker than that required over the region where oxygen ions dominate. The downward force on an oxygen ion is now larger, and the concentration of oxygen ions decreases more rapidly with altitude than before.

The dotted lines in Figure 2.1 represent the distribution of each constituent over a region of chemical or photochemical equilibrium. It may be seen that the altitude dependence of the hydrogen ion concentration in the region of chemical equilibrium is the same as that of diffusive equilibrium. There is no physical reason for this; it simply happens that the distribution given by Equation (2.4) is the same as that obtained from Equation (2.25) when it is assumed that H^+ ions are a minor constituent in a mixture of ions of He^+ and O^+.

Inspection of Equation (2.4) will reveal that once a model for the neutral atmosphere has been chosen, the concentration of H^+ ions over the region of chemical equilibrium is fixed by the distribution of O^+ ions over this region. This latter distribution is easily inferred from the altitude and concentration of the F2-layer maximum, which is reasonably well known as a function of local time and epoch of the solar cycle. Thus, the specification of a model topside consisting only of ions of H^+ and O^+ at a particular time depends only on the choice of the ratio n(H)/n(O). The amount of He^+ ions present in the topside is, however, in no way dependent upon the amount of ionized hydrogen and oxygen present, and the He^+ ion concentration at, say, 500 km may therefore be regarded as a disposable parameter.

Within this context it is interesting to examine the effect on the structure of the topside of varying the helium ion content. The broken lines in Figure 2.1 represent the distributions that ions of O^+ and H^+ would take up in the absence of He^+ ions. The difference is noticable only over a rather restricted range of altitudes near 1500 km. The important conclusion illustrated here is that, under conditions of diffusive equilibrium, the concentration of H^+ ions at high altitudes--and most particularly in the equatorial

plane--is determined by the concentration of these ions in the region of chemical equilibrium and is independent of the amount of ionized helium present in the topside. Further reference to this statement will be made in the next section.

2.5 The Structure of the Protonosphere

Although the helium ion content of the topside ionosphere is somewhat uncertain, it is probable that hydrogen ions become the dominant constituent somewhere in the region between 1000 and 2000 km above the surface of the earth at middle geomagnetic latitudes. As seen from Figure 2.1 and the discussion of the preceding section, this transition is very gradual. Above the transition level, the concentration of hydrogen ions decreases with altitude, with the result that very low concentrations (of the order of 10^3 cm^{-3}) are found in the protonosphere.

In deriving the hydrostatic equations which govern the distribution of the ions and electrons in the topside--Equations (2.16) and (2.17)--the implicit assumption was made that the ionization behaves as a continuous fluid. The same assumption is also made in calculating the distribution adopted by the neutral constituents of the atmosphere. In the case of the neutral atmosphere it is known that this assumption is no longer valid at altitudes above about 500 km, where the mean free path of the particles is sufficiently large that most of them execute ballistic trajectories under the influence of gravity. This region is known as the exosphere. The distribution of the atmosphere in this region may be found, under equilibrium conditions, from Liouville's theorem. If the possibility of escape of particles is excluded, the resulting distribution is the same as the hydrostatic distribution. Escape of particles

does occur, however, introducing an anistropy in the velocity distribution of the exospheric particles. The result is a distribution that departs from one of hydrostatic equilibrium.

The justification for treating the ionized component of the atmosphere as a fluid, even in the protonosphere, is that escape of charged particles is prevented by the geomagnetic field. A complete description of the ionization in the protonosphere is complicated by the presence of high-energy trapped particles. The concentration of these particles is, however, very much less than that of the ionization at thermal energies and does not influence the distribution of this ionization. It should also be noted that the cross section for coulomb interaction is, for the temperature characteristic of the protonosphere, sufficiently large that the mean free path of both electrons and ions is much less than the distance between conjugate ends of the magnetic field lines.

Thus the distribution of ionization throughout the protonosphere is governed by a hydrostatic relationship. This distribution is the same for both electrons and hydrogen ions, except for a difference near the base of the protonosphere that is due to the presence of helium ions.

Because of the great extent of the protonosphere and the curvature of the lines of force of the geomagnetic field over this region, it is necessary to work in terms of geocentric distance r rather than altitude. When allowance is made for the equality of electron and ion temperatures in the protonosphere and the fact that the mean ionic mass is simply that of a hydrogen ion, Equation (2.20) becomes

$$\frac{dn_e}{dr} = - n_e\left(\frac{m_1 g}{2kT_e}\right), \qquad (2.26)$$

where m_i is the hydrogen ion mass. This is the hydrostatic equation for a gas of particles of mass $m_i/2$ in an isothermal state. The quantity

$$H_i \quad \frac{2kT_e}{m_i g} \qquad (2.27)$$

has the dimensions of length and is known as the scale height of the ionization. These equations are the basis for the treatment to follow in the remainder of this chapter. For the purpose of simplifying the notation, the subscripts e and i will be supressed. No ambiguity arises, as the temperature and concentration of the electrons and ions are equal.

The ionization in the protonosphere is constrained by the geomagnetic field to co-rotate with the earth (Johnson, 1960, Hines, 1963). A centrifugal term should therefore be added to the right-hand side of Equation (2.26). This term becomes appreciable only at distances in excess of 5 earth radii. As the lines of force of the geomagnetic field do not rise to such heights at middle latitudes, this term will be neglected. The expression for the electron density distribution is thus obtained by integration of Equation (2.26), which gives

$$n \quad n_o \exp\left[\frac{r_o^2}{H_o}\left(\frac{1}{r} - \frac{1}{r_o}\right)\right], \qquad (2.28)$$

where the subscript zero is used to denote quantities evaluated at the base of the protonosphere.

The ionization in the protonosphere is further constrained by the geomagnetic field to move only along the lines of force. It was first pointed out by Johnson (1960) that, as a consequence of this, the distribution of ionization

is hydrostatic along a line of magnetic force rather than in the vertical. The distribution is still given by Equation (2.28) provided that the tip of the radius vector r is taken to lie always on the same line of force. This requirement is satisfied if the equation of the lines of force of a dipole field,

$$\frac{r}{r_o} = \frac{\cos^2\theta}{\cos^2\theta_o} \quad , \tag{2.29}$$

is used to eliminate r as the variable in Equation (2.28), to give

$$n = n_o \exp\left[\frac{r_o \cos^2\theta_o}{H_o} \sec^2\theta\right] \exp\left[-\frac{r_o}{H_o}\right] . \tag{2.30}$$

This is the equation for the variation of electron density along the line of force which passes through the point (r_o, θ_o). Since the quantity inferred from propagation of atmospheric whistlers is electron density in the plane of the geomagnetic equator, the relation given by Equation (2.30) is more useful if expressed in terms of this quantity. It follows from Equation (2.30) that the electron density in the equatorial plane n_t is given by

$$n_t = n_o \exp\left[\frac{r_o \cos^2\theta}{H_o}\right] \exp\left[-\frac{r_o}{H_o}\right] . \tag{2.31}$$

When this equation is solved for n_o and substitution into Equation (2.30) is made the expression

$$n = n_t \exp\left[\frac{r_o \cos^2\theta_o}{H_o} \tan^2\theta\right] , \tag{2.32}$$

is obtained.

The distribution of ionization within the protonosphere that results for a temperature of 1250°K when values of n_t are taken from the electron density profile in the equatorial plane inferred from whistler data is shown in Figure 2.2. This figure is taken from Johnson (1962); the electron density in the equatorial plane is based on average whistler data (Smith, 1961). The most striking feature of this figure is the strong latitude variation deduced for the electron density at the base of the protonosphere. According to this result, the electron density at the 1000 km level should decrease by an order of magnitude between latitude 45° and latitude 65°. As Johnson (1962) has pointed out, this is difficult to accept in view of the relative constancy of the density at the F2-layer maximum over this range of latitude.

That this result does not necessarily imply a latitude variation of electron density at the 1000 km level became evident with the realization that helium ions are present in sufficient quantity in the topside ionosphere to be the major constituent at this altitude. However, this does not remove the difficulty, since, apart from a relatively small correction for the presence of helium ions in the lower part of the protonosphere, the curves shown in Figure 2.2 represent the distribution of hydrogen ions. Thus, acceptance of the whistler data and the concept of diffusive equilibrium along magnetic field lines leads to the conclusion that the concentration of hydrogen ions at the 1000 km level decreases by nearly an order of magnitude over the range of latitude from 45° to 65°.

Recent measurements of electron density over the 1000 km surface by the Alouette satellite indicate a decrease of electron density with latitude (Thomas and Sader, 1964). Within the latitude range of 45° to 65° this decrease is very much less than the decrease in the hydrogen ion concentration shown in

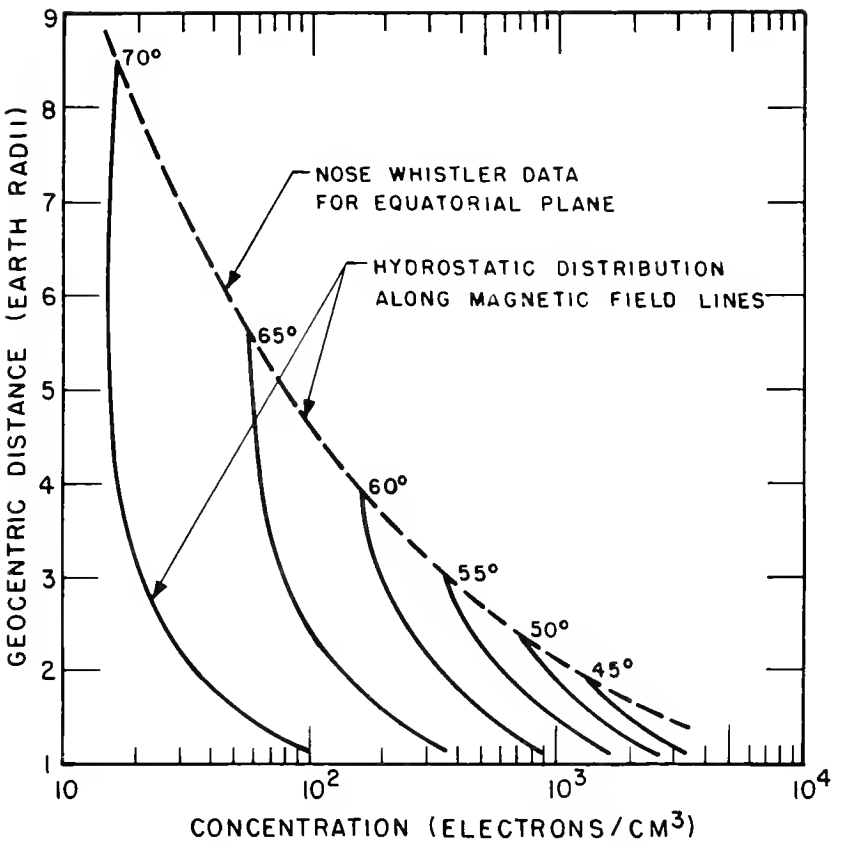

Figure 2.2 Distribution of ionization along magnetic field lines leaving the earth at various latitudes (after Johnson, 1962). The broken line represents the distribution in the equatorial plane, inferred from nose whistler data.

Figure 2.2. If the decrease of hydrogen ion concentration with latitude is as large as indicated here, the only information which these measurements provide on the ionized hydrogen present in the topside is that it is a minor constituent at the 1000 km level within this range of latitudes.

The possibility that the method used to infer the latitude dependence of hydrogen ions may be in error should not be overlooked. The distribution is assumed hydrostatic along the field lines for a latitude-independent temperature of 1250°. It will be shown in Chapter 6 that the temperature of the protonosphere is probably at least a factor of 2 greater than this. However, increasing the temperature serves only to increase the scale height. This has the effect of straightening out the curves in Figure 2.2, bringing the concentration at 1000 km more in line with that in the equatorial plane, and thus increasing the latitude variation. With regard to the latitude variation of temperature, it should be remarked that recent measurements by the Ariel satellite do indicate a dependence of electron temperature upon latitude (Willmore et al., 1964). This dependence is, however, confined mainly to latitudes below 30°. What little variation is shown at higher latitudes is such that the electron temperature increases with latitude. This increases the scale height at the higher latitudes over that at lower latitudes, and again, the result is an increase in the latitude variation of hydrogen ion concentration.

It may be fairly stated that a satisfactory explanation of the latitude variation of hydrogen ion concentration at the 1000 km level has not yet emerged. Whether there is in fact a variation of the magnitude presently indicated is a question that will probably not be resolved until extensive measurements of the hydrogen ion concentration over the 1000 km surface become available.

2.6 Measurements in the Topside Ionosphere and the Protonosphere

2.6.1 Langmuir Probes

If a small metallic sphere is placed in a plasma it is observed that the sphere takes up a potential that is negative with respect to that of the plasma. This is a consequence of the fact that the mean thermal velocity of the electrons is greater than that of the ions and more electrons than ions strike the surface of the sphere per unit time. If now the sphere is in some manner grounded to plasma potential, a current will flow to the sphere. With the further addition of a source of emf in the circuit to provide a variable retarding potential on the sphere, the apparatus becomes a simple form of Langmuir probe. It is easily shown that the electron current to such a probe in the retarding potential region ($V \leq 0$) is given by

$$I_e = \left(\frac{eAn_e \bar{v}_e}{4} \right) \exp \left[\frac{eV}{kT} \right], \qquad (2.33)$$

where A is the surface area, e is the magnitude of the electronic charge, $\bar{v}_e$ is the mean thermal velocity and V is the potential of the probe relative to plasma potential. This expression is independent of the geometry of the probe.

The probe is operated at a voltage which is repeatedly swept from a negative value, that appreciably retards the electrons, to a small positive value. Inspection of Equation (2.33) shows that the electron temperature may be determined from the current-voltage characteristic of the probe independent of electron density and surface area. Once this has been done, the electron density may be determined from the observed current at plasma potential. In practice, the potential of the sphere relative to plasma potential cannot be measured directly,

but the point on the current-voltage characteristic at which this condition obtains may be identified.

Rocket-borne Langmuir probes have been used extensively for measurements in the lower ionosphere. With the exception of electron temperature profiles in the F region,--to be enlisted in support of the theory presented in Chapter 5--these measurements are not relevant to the structure of the topside and will not be discussed here. Measurements of electron temperature and density over the topside ionosphere have been obtained from modified Langmuir probes flown on the satellites Explorer VIII and Ariel I. The electron temperature measurements obtained by Ariel I are of particular interest and will be described here.

The satellite went into operation ten days after launch on April 26, 1962. It functioned well until it was damaged by the high-altitude nuclear explosion of July 9; intermittent data were received thereafter until August 22. Some 70,000 measurements of density and temperature were made during the period. The type of Langmuir probe used was characterized by an applied voltage which, in addition to the sweep voltage, possessed a component made up of two small-amplitude sine waves of similar frequency (the Druyvesteyn probe). Certain advantages in telemetry accrued from operation of the probe in this mode, it was also observed that the effect of the photoemission from the probe was eliminated, as were variations in the positive ion current to the probe resulting from satellite rotation.

The results of the reduction of the electron temperature data have been published recently (Willmore et al. 1964). The important features are the following.

(i) Diurnal variation

The electron temperature at 1000 km level increases rather

rapidly from a minimum value of about 1200°K near 0400 hours local time, leveling off near 0800 hours at about 2100°K. A decrease sets in near 1600 hours local time and this continues at a uniform rate until 2200 hours, with a slower decrease thereafter. No important differences in this pattern are found over the range of altitudes between 600 km and 1200 km.

(ii) Altitude variation

The region between 600 km and 1200 km is very nearly isothermal. A small positive (temperature increasing with altitude) gradient is found. There is no significant difference in the magnitude of the gradient between day and night. Limited data available for the 400 km level indicate that the gradient increases with decreasing altitude.

(iii) Latitude variation

The electron temperature is found to increase with geomagnetic latitude, but this is confined mainly to low latitudes. At the 400 km level the noon temperature increases from 1000°K at the equator to about 1500°K at latitude 40°; no further increase is observed at this level between latitudes of 40° and 60°.

The observational evidence summarized here will be cited in Chapter 5, where theoretical electron temperature profiles are presented, and again in Chapter 6 in connection with the predicted degree of nighttime thermal coupling between the ionosphere and the protonosphere.

2.6.2 Ion Mass Spectrometers

If the range of the sweep voltage is adjusted to include larger positive values, the Druyvesteyn probe may be employed to measure ion density and

temperature. Such a probe, encased in a grid biased to eliminate a large electron current, was also flown on the Ariel I satellite. The reduced data gives the energy spectrum of each ion species present. As seen from the satellite moving with velocity v_s, the energy of an ion at rest in the ionosphere is $m_i v_s^2/2$. As a consequence of the fact that v_s is considerably in excess of the mean thermal velocity of even the lightest ionic constituent present, the energy distributions of the constituents appear separated. The probe thus serves as a mass spectrometer.

The three peaks corresponding to ions of O^+, He^+ and H^+ have all been detected in the reduced data. These measurements have been used to infer the transition levels where ions of He^+ and ions of H^+ first become the dominant constituent by assuming a diffusive equilibrium model of the topside with temperature equal to that of the neutral atmosphere (Bowen et al., 1964). The results may be briefly summarized as follows:

(i) Helium ions

The helium ion transition altitude over the range of geomagnetic latitude 35^o to 45^o is characterized by a strong diurnal variation: it is found near 600 km at midnight and rises to 1200 km during the day; there is marked asymmetry in that the maximum altitude is not attained until 1700 hours local time.

(ii) Hydrogen ions

During the day, hydrogen ions were in such a minority below 1200 km that a good estimate of their concentration could not be obtained by this technique. Some measurements were obtained for late night and early morning. These indicate a hydrogen ion transition level near 1600 km.

The helium ion content of the topside inferred from a measurment by the Explorer VIII satellite is somewhat greater than that indicated by the Ariel I measurements: a single measurement at 0615 local time on November 24, 1960 (Bourdeau et al., 1962) indicates a transition to He^+ ions at about 1600 km. This is nearly 1000 km higher than the transition altitude deduced from the Ariel I data for this time. It is unlikely that the difference of temperature between late fall of 1960 and summer of 1962 could have been large enough to bring about a change of transition altitude of this magnitude.

The ion mass spectrometer flown in Explorer VIII is one of a class of ion traps, so-called because the ion collector is not fully exposed. This particular type consists of a collector located at the base of a slot in the satellite body. A variable retarding potential is applied to the collector. This field is kept inside the satellite by a grid located over the aperture of the slot. A negatively biased grid situated just above the collector serves the dual purpose of repelling the incoming electrons and suppressing photo-emission from the collector. In general, ion traps have better resolution than does the Druyvesteyn probe when used as mass spectrometer. Results from Explorer VIII were rather limited, however. In particular, no measurements were made of the concentration of H^+ ions due to an unfavorable orientation of the trap which seems to have persisted at altitudes above 1600 km.

Rocket-borne ion traps have made a notable contribution to the knowledge of the composition of the topside. Although employed here only as ion collectors rather than mass spectrometers, they are able to provide useful information on composition through detection of the change in the total ion distribution that occurs at a composition transition level. It has already been noted that the

presence of helium ions was first detected from a measurement of total ion concentration by Hale (1961). More recently, a rocket-borne ion trap has been used to detect a nighttime transition level for He^+ ions near 600 km (Donley, 1963)--a figure consistent with the Ariel I deductions.

Measurements obtained with a radio frequency mass spectrometer of very high resolution have been reported by Taylor et al., (1963). This technique provides complete profiles of helium and hydrogen ion concentration from 350 km to 950 km. The shape of the hydrogen ion profile has been used by Hanson et al., (1963) to deduce both the rate coefficient, K, for the charge exchange process and the magnitude of the diffusive flow of hydrogen ions at the time of the measurement.

2.6.3 The Topside Sounder Alouette

The Alouette satellite was placed into a nearly circular polar orbit on 29 September 1962. Mean altitude of the orbit is about 1000 km. The frequency of the sounding signal is swept over the range from 0.45 to 11.8 Mc/s, which is roughly the range of plasma frequency between 1000 km and the F2-layer maximum. The frequency of the signal and the time lapse between transmission and reception of the reflected signal are telemetered to the ground. A complete ionogram--as the plot of frequency versus delay time is called--is available every 18 seconds, during which time the satellite travels about 120 km. During the first 8 months of operation some 250,000 ionograms were thus received and stored at various telemetry stations, and this rate of accumulation of data continues at the present time. Only a very small fraction of these ionograms have been reduced to obtain topside electron density profiles. Two systematic programs of data reduction, the results of which are of interest here, are the projects at Stanford (Thomas and Sader, 1964) and at Slough, England (King et al., 1963a,b).

Observations made at Stanford have been used to obtain electron densities at the altitude of the satellite orbit. The decrease of electron density with latitude along the 1000 km surface thus deduced has already been cited in the discussion of the distribution of ionized hydrogen. A further result which has come from this analysis is that the daytime electron density along the 1000 km surface at middle geomagnetic latitudes is nearly a factor of two lower in winter than in summer. The daytime electron density at the F2-layer maximum is known to behave in the opposite way, however, being higher in winter than in summer. Taken together, these observations imply the existence of a seasonal variation in either the temperature or the composition of the topside, or both.

A large number of electron density profiles have been obtained from reduction of ionograms at Slough. Most of these pertain to low latitudes and the interpretation given them is not of immediate concern here. One important conclusion that has been drawn from the interpretation of mid-latitude profiles is that the electron temperature increases with latitude. Other conclusions are that the altitude of transition from oxygen ions to helium ions during the day is consistent with that inferred from the Ariel I measurements and that a gradient of either T_i or T_i/T_e is present over part of the topside.

The information summarized here is descriptive of the gross features of the topside and is therefore of the type which a first order theory of the region, such as that described in this work, should be able to interpret. As this section is a review of topside measurements, however, it should be mentioned that the Alouette satellite has also provided a valuable insight into what may be termed the 'micro-structure' of the topside: columns and sheets of enhanced ionization oriented along the geomagnetic field, some of which are very short-lived and others of which appear to be permanent features. Then too, the

ionograms exhibit a number of hitherto unsuspected features that have stimulated a considerable amount of research on the subject of transmission of electromagnetic waves from an antenna immersed in a plasma.

2.6.4 Incoherent Backscatter

Conventional sweep frequency ionospheric sounding depends upon the fact that the ordinary circularly polarized electromagnetic wave is reflected at a point in the ionosphere where the plasma frequency is equal to that of the incident wave. Thus, a wave of frequency greater than that of the plasma frequency at the F2-layer maximum penetrates the ionosphere and continues into interplanetary space. A very small fraction of the power of such a signal should, however, be returned by means of incoherent scattering by electrons throughout the ionosphere. It was first proposed by Gordon (1958) that, given a transmitted signal of sufficient power, the fraction thus returned may be detectable. A considerable amount of theoretical and experimental work has followed, including the investment of large sums into the design and construction of high-powered radar equipment.

From a simple physical viewpoint this technique has obvious attractions. Since the fraction of the transmitted power that is scattered is a function of the electron density, all that is required to obtain an electron density profile is a measurement of the strength of the return signal as a function of time. The altitude range of such a profile would be limited by the sensitivity of the receiving equipment rather than the height of the F2-layer maximum. Moreover the signal returned from each level would be expected to show Doppler broadening due to thermal motions of the electrons, thus providing information from which an electron temperature profile could be derived. It is now known

that the situation is far more complex than this. The electrons and ions are coupled by coulomb forces, and the response of the electrons to the incident wave is largely controlled by the ions. Thus, the scattering by the electrons is not incoherent. This has greatly complicated the theoretical analysis required to relate the measurements to the physical quantities. But with this has come the realization that information on the ion temperature as well as the electron temperature and electron density is contained in the spectrum of the signal returned from a given level.

A good accound of the development of this field has been given recently by Evans and Loewenthal (1964). The same paper also describes in detail the equipment and method of data reduction used at the Millstone Radar Observatory located near Lexington, Massachusetts. As reference to temperature profiles deduced from measurements made at this station will be made in Chapter 5, it is pertinent to summarize briefly factors which influence the accuracy of these measurements and their reduction.

Observations are limited by ground clutter to altitudes above 200 km. An upper limit of about 750 km is imposed by the weakness of the returned signal. It is possible to deduce electron density, electron temperature and ion temperature over this altitude range. The largest source of error arises from the fact that interpretation of the spectra requires a knowledge of the relative abundance of the ion species present. In practice, a composition of pure oxygen ions is assumed, but the temperature profiles deduced are altered considerably by the introduction of a relatively small percentage of helium ions.

2.6.5 Atmospheric Whistlers

A lightning discharge gives rise to a broad spectrum of electromagnetic radiation. The very low frequency components of this pulse can penetrate the F2-layer and propagate thereafter along a line of force of the geomagnetic field, across the equatorial plane and into the conjugate hemisphere. The protonosphere behaves as a highly dispersive medium for this mode of propagation, and the higher frequencies arrive first at the conjugate end of the field line. If an audio amplifier is connected directly to an antenna in this vicinity, a tone which begins at a high pitch and descends rapidly to a low one is heard. Quite often, an appreciable fraction of the energy of the signal is reflected in the conjugate hemisphere, and a short time later a whistler may be detected in the region of the original lightning flash.

A discussion of the way in which information on the structure of the protonosphere may be extracted from this phenomenon is contained in a recent review by Carpenter and Smith (1964). The observable property of a whistler is the frequency of the tone as a function of time. An example of a frequency versus delay time curve is shown in Figure 2.3. The origin of the time scale can in practice be identified by the time of arrival of high frequency radiation direct from the lightning stroke. The existence of a minimum delay time (appearing in this figure near 5 Kc/s) is characteristic only of whistlers which propagate at geomagnetic latitudes in excess of about 45°. Whistlers which show this effect are referred to as nose whistlers.

The delay time as a function of frequency is given by

$$t = \frac{1}{2c} \int_0^s \frac{f_m f_h}{f^{1/2} (f_m - f)^{3/2}} \, ds \, , \qquad (2.34)$$

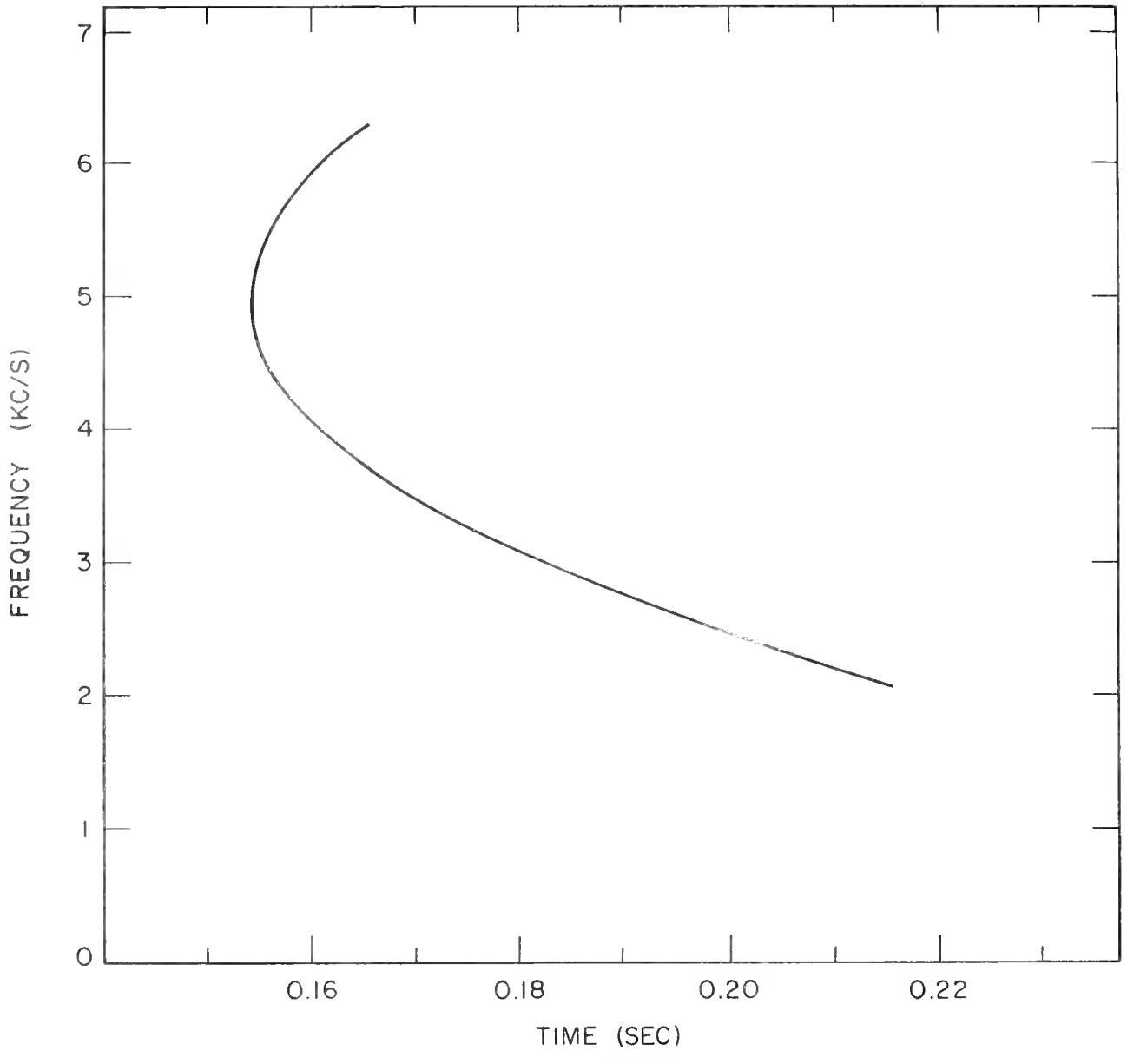

Figure 2.3 Frequency versus delay time for a whistler recorded at 1235 on March 19, 1959 at College, Alaska (after Pope, 1959).

where f is the frequency, f_n and f_h are, respectively, the local plasma frequency and electron gyrofrequency, and s is the distance traveled along the field-aligned path. For an assumed electron density distribution along the path, it is possible to integrate Equation (2.34) to obtain a theoretical frequency versus delay time curve; the model electron density distribution giving the closest approximation to the observed curves at all latitudes corresponds most closely to the actual distribution within the protonosphere.

In fact, the integral in Equation (2.34) is rather insensitive to the distribution over much of the field-aligned path, and a wide variety of models give the same result. The integral is, however, sensitive to the magnitude of the electron density near the top of the path, and it is therefore possible to select a model for the altitude variation of electron density within the plane of the geomagnetic equator. One such model, due to Smith (1961), appears in Figure 2.2.

At frequencies below that of the frequency of minimum delay the relation of frequency to delay time is such that $tf^{1/2}$ is approximately constant. This quantity is called the dispersion of a whistler (Storey, 1953). Since at these frequencies it is also true in general that $f \ll f_h$ at all points along the path, the dispersion, D, is given by

$$D = \frac{1}{2c} \int_0^s \frac{f_n}{f_h^{1/2}} \, ds \, . \qquad (2.35)$$

By definition, the dispersion is also a sensitive function of electron density at the top of the whistler propagation path. Thus, the electron density distribution in the equatorial plane has also been inferred from the latitude variation of this quantity (Alicock, 1959).

The dependence of the dispersion upon electron density at the top of the propagation path is of importance in still another respect; namely, that at a given latitude the diurnal variation of the whistler dispersion is a measure of the corresponding variation of electron density in the field tube along which the whistler propagates. There is some evidence that the dispersion exhibits a diurnal variation of 20 or 30 per cent at middle geomagnetic latitudes (Outsu and Iwai, 1959; Rivault and Corcuff, 1960), though such behavior is by no means consistent, there being occasions where no significant variation has been detected (Carpenter, 1962).

A variation of electron density at the top of the field tube may be the result of a changing distribution within the tube due to temperature changes, or a change in the total content of the tube, or both. It is the purpose of the section following to report the effect on the whistler dispersion of the redistribution of electrons resulting from a change of temperature in an isothermal protonosphere.

2.7 The Relation Between the Dispersion of a Whistler and the Electron Temperature in the Protonosphere

An increase in the temperature of the protonosphere will increase the scale height of the ionization, thus tending to make the distribution more nearly uniform along the field tube. If it is assumed that no ionization enters or leaves the tube as it is heated, the density at the equatorial face of the tube will increase while that at the base will decrease. The effect of this redistribution on the whistler dispersion may be calculated from Equation (2.35) when an expression for the electron density as a function of temperature has been obtained. This expression is given by Equation (2.32)

when the explicit dependence of n_t upon temperature is known. As n_t is related to both the temperature and the total content of the field tube, an expression for the temperature dependence of n_t follows from integration of the electron density over the volume of the field tube. In the coordinates r and θ, the total electron content of a field tube bounded by the equatorial plane at one end and intersecting unit area on a sphere of radius r_o and latitude θ_o is given by

$$N(\theta_o) = 2 n_t r_o \frac{\tan\theta_o}{\cos^7\theta_o} I(\theta_o, T) \quad , \tag{2.36}$$

where

$$I(\theta_o, T) = \int_o^{\theta_o} \exp\left[\frac{r_o \cos^2\theta_o}{H_o} \tan^2\theta\right] \cos^7\theta \, d\theta,$$

and T is the common electron-ion temperature, entering the integrand through the temperature dependence of the scale height.

A change in the protonosphere temperature will alter the distribution of ionization within the field tube in a manner that can be calculated from this equation. Assuming that the total electron content of the field tube remains constant as the temperature goes from T' to T, the electron density at the equatorial face of the tube is given by

$$n_t(T) = n_t(T') \frac{I(\theta_o, T')}{I(\theta_o, T)} \quad . \tag{2.37}$$

The integral $I(\theta_o, T)$ has been evaluated numerically for a number of different temperatures for each of two latitudes: $\theta_o = 40^o$ and $\theta_o = 60^o$. The variation

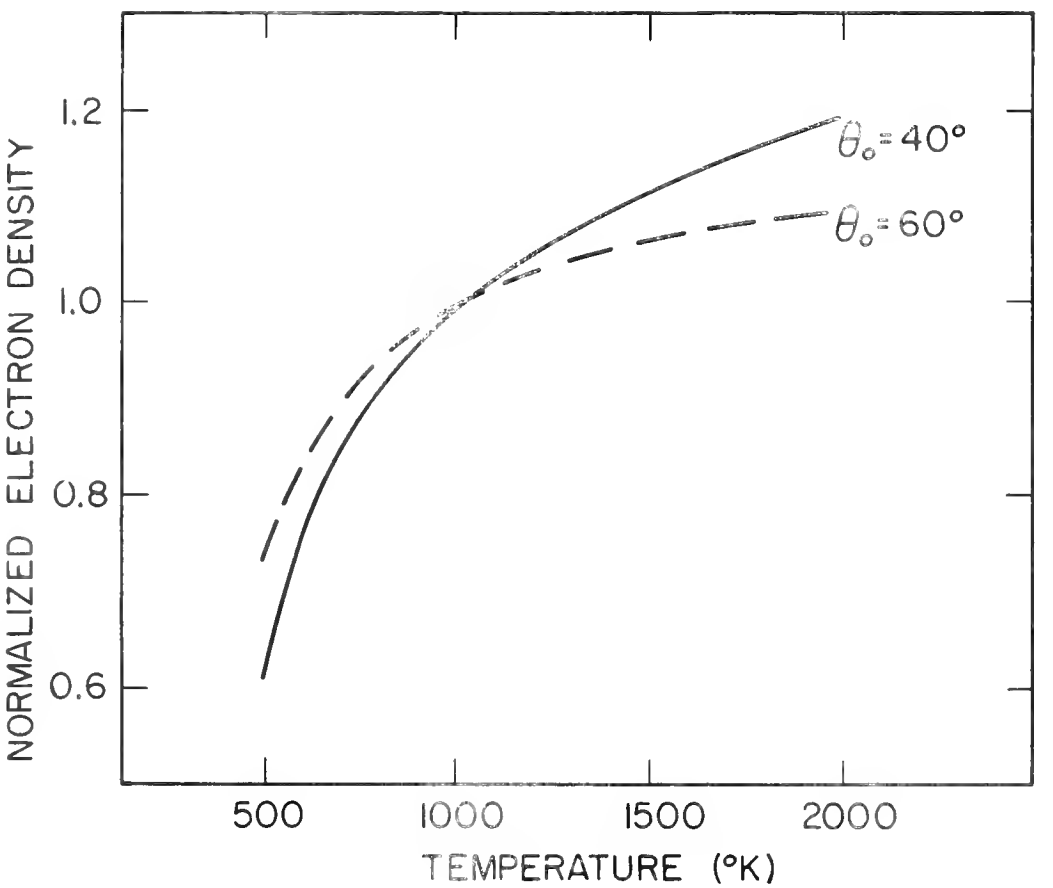

Figure 2.4 The variation with temperature of the electron density at the equatorial face of the field tube intersecting unit horizontal area 1000 km above the earth's surface at geomagnetic latitude θ_o. The total electron content of the field tube is assumed to remain constant.

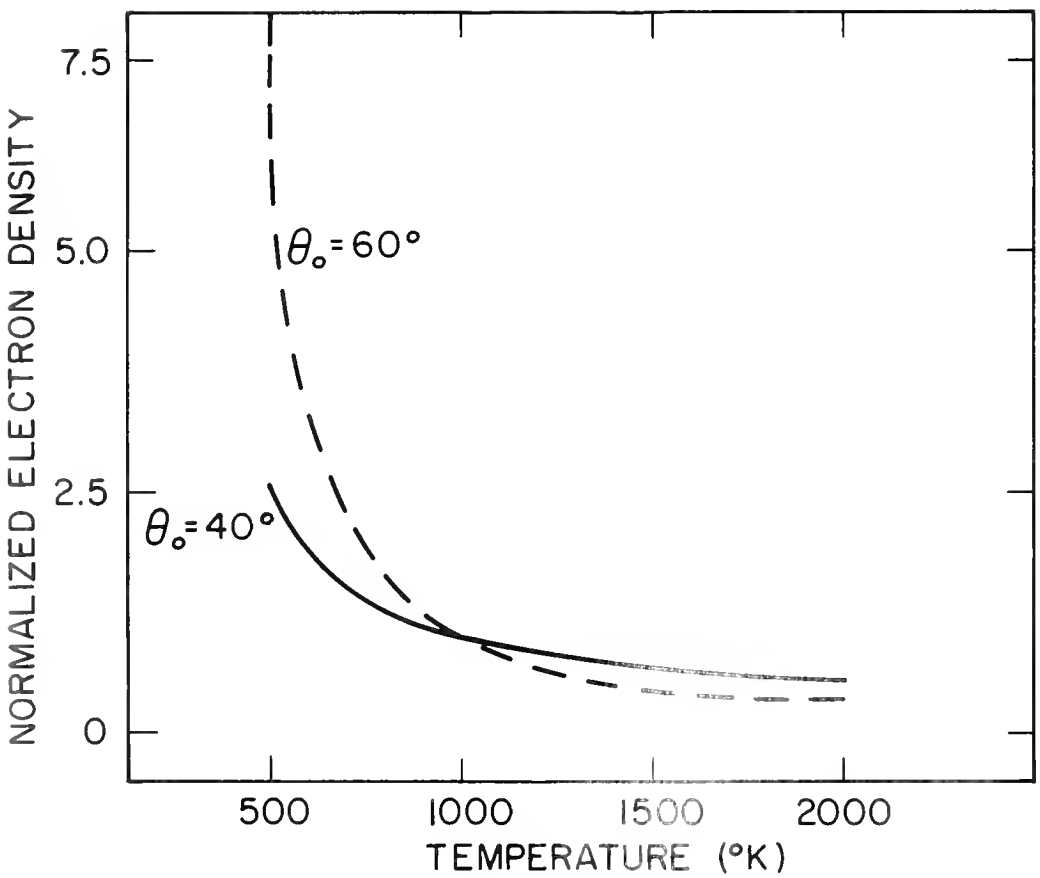

Figure 2.5 The variation with temperature of the electron density at the lower boundary of the field tube intersecting unit horizontal area 1000 km above the earth's surface at geomagnetic latitude θ_o. The total electron content of the field tube is assumed to remain constant.

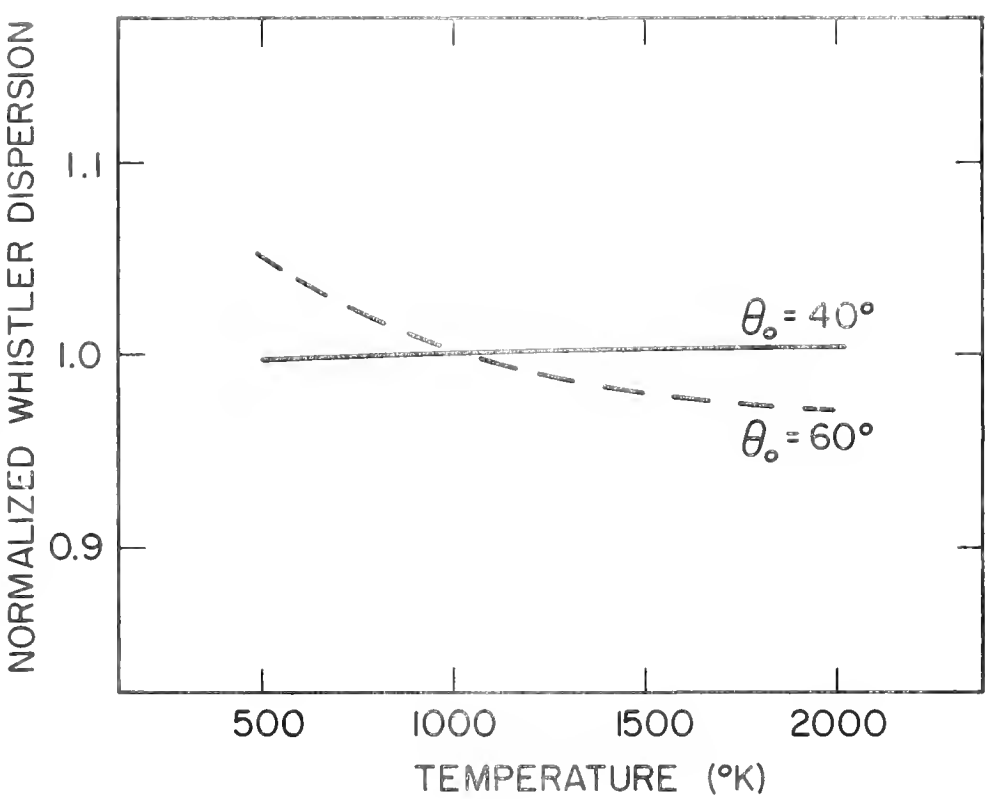

Figure 2.6 The calculated dependence of the whistler dispersion upon protonosphere temperature for propagation path with lower boundary at latitude θ_o. The dispersion has been normalized to unity for a temperature of $1000^{o}K$.

thwell (1962) pointed out that the diffusion time of protons through atomic
ygen ions, being impeded by coulomb forces, is highly temperature dependent,
d might therefore be expected to proceed much more rapidly during the day
an at night. She therefore suggested that the decrease in summer hemisphere
-layer electron density, and the increase in electron density in the opposite
misphere, might be due to diffusion of ionization along lines of the geomag-
tic field between conjugate points in the two hemispheres. This explanation
ll be stated here in the form of a hypothesis to be tested by the steady-
ate theory of ionosphere-protonosphere coupling.

With the cessation of production of ionization by solar radiation at sunset,
e F2-layer begins to decay. Two modes of decay are possible, depending upon
e relative importance of diffusive transport. The location of the maximum
te of loss is a short distance below the layer maximum. If downward transport
ionization by diffusion is negligible, the layer decays from the bottom up-
rd. This results in an increase of the altitude of the layer maximum with
me during the night. If, on the other hand, diffusive transport is efficient,
nization diffuses downward to fill up the gap in the region of maximum rate
decay. This results in a layer which decays uniformly, the altitude of the
ximum remaining relatively constant with time throughout the night. The
rmer mode of decay seems to be most closely in agreement with observation,
rticularly during the post-midnight hours. However, the diffusion rate re-
ired for this mode of decay is much slower than that expected from theore-
cal considerations. In view of this discrepancy, it is necessary to consider
e possible influence of phenomena other than diffusion and loss.

These phenomena may be divided into two categories: those which hinder
e diffusion of ionization downward to a region of high loss rate, those which

replenish ionization. Insofar as they are responsible for an upward drift of ionization, atmospheric winds and electrodynamic forces belong in the first category. Replenishment of ionization may result from the existence of a flux of ionizing corpuscular radiation or from a source of ionization external to the layer.

The hypothesis to be adopted here is that the nighttime F2-layer is maintained by a flow of ionization into the top of the layer. The analysis of Hanson and Ortenburger (1961) shows that the coupling between the F2-layer and the protonosphere is weak at night, in the sense that only a small fraction of the ionization in a protonospheric field tube diffuses downward through the charge exchange region and into the F2-layer. However, the content of the protonosphere is sufficiently large that the possibility that the downward flow may be enough to maintain the nighttime F2-layer cannot be ruled out.

3.1.3 Purpose of the Chapter

The coupling between the F2-layer and the protonosphere is regulated by the rate at which hydrogen ions can diffuse through the region between the base of the protonosphere and the chemical equilibrium region for charge exchange with oxygen ions. The analysis of this process forms the subject matter of the next chapter. The purpose of this chapter is to investigate the effect of an external source or sink of ionization upon the structure of the F2-layer, and, in particular, to obtain an estimate of the magnitude of the influx or efflux of ionization which is required to produce a significant change in the structure of the layer.

The equations which govern the altitude dependence of the rate of production, loss and diffusive transport are presented in Section 3.2. The diffusion

in Equation (2.38). For a given field tube

$$D \propto N^{1/2}(\theta_o) \quad .$$

For the value of the critical flux quoted above, the expected diurnal variation of D at latitude 40° would be around 5 per cent, while for the increased value, the variation would approach 15 per cent.

The magnitude of the dispersion may be calculated for a given field tube content from Equations (2.36) and (2.38). For a whistler which originates in the conjugate hemisphere, the total field tube content is $2N(\theta_o)$ and the value of D obtained from Equation (2.38) must therefore be multiplied by a factor of $\sqrt{2}$. This calculation gives $D = 57$ sec$^{1/2}$ for a whistler which propagates between hemispheres along the field line intersecting the 1000° km surface at latitude 40°, when $N(40^{\circ}) = 4.52 \times 10^{12}$ cm^{-2}. This may be compared with the figure of 70 sec$^{1/2}$ taken from the curve of dispersion versus latitude given by Allcock (1959) and based on whistler data recorded at a number of stations over the period January to May 1957. The value of D derived here is within the limits of experimental error for these data.

3. THE STEADY-STATE THEORY OF THE STRUCTURE OF THE F2-LAYER

3.1 Introduction

The F2-layer is composed of a plasma of atomic oxygen ions and electrons. At the maximum of the layer, the electron density is typically about two orders of magnitude lower than the concentration of the dominant constituent of the neutral atmosphere in this region, atomic oxygen. The distribution of ionization is governed by the relative importance of the competing processes of production, loss and diffusion.

3.1.1 Discussion of Physical Processes

The plasma of atomic oxygen ions and electrons is produced by photoionization of atomic oxygen at a rate per unit volume which attains a maximum value at the peak of the F1-layer (about 180 km), a condition which is in fact responsible for the existence of this inconspicuous layer. The region at and above the peak of the F1-layer is one of photochemical equilibrium. The electron density increases with altitude above this peak because the rate of loss of ionization decreases more rapidly with altitude than does the rate of production.

The increase of ion and electron concentration with altitude is attended by a downward diffusion of this ionization through the neutral atmosphere. The magnitude of the diffusive flow increases with altitude for a given gradient of ion concentration, due to the decrease of the concentration of the neutral atmosphere. On the other hand, the downward flow through any level can be maintained only at the expense of the ionization produced above that level. Since the rate of production decreases with altitude, it is evident that the increase of electron and ion concentration with altitude cannot continue

indefinitely. A departure from photochemical equilibrium must ultimately take place, with the distribution of ionization approaching the exponentially decreasing distribution of diffusive equilibrium at great heights. It follows from these considerations that the distribution of ionization above the F1-layer attains a maximum.

The competing processes of production, loss and movement of ionization are related through the continuity equation. With suitable assumptions regarding the structure of the neutral atmosphere it is possible to obtain an electron density profile for the F2-layer by solving the continuity equation. Several approaches to this problem have been tried: analytical solution for a particular case (Yonezawa, 1956, 1958) and for a more general case (Bowhill, 1962a, 1962b), solution by analog computer (Rishbeth, 1963, 1964); and numerical solution by digital computer (Rishbeth and Barron, 1960; Gliddon and Kendall, 1960). Although this theory gives a layer of the shape and density of the F2-layer, it has not been entirely successful in interpreting the behavior of the layer, even when the time-dependence of the physical quantities is taken explicitly into account. Apparently, this is due to the fact that processes other than production, loss and diffusion occur and at times are dominant. The effects which have been most intensively investigated are those which arise from movement of ionization by either electromagnetic forces or atmospheric motions. Efforts along this line have been hampered not only by the mathematical difficulty which such considerations lead to, but also by lack of quantitative knowledge of these phenomena.

Still another factor which may influence the behavior of the layer is the existence of a source or sink of ionization situated above, and external to,

e F2-layer (Hanson and Ortenburger, 196, Bothwell, 1962). Such a region is
e protonosphere, the exchange of ionization being effected by the chemical
rocess of charge exchange involving oxygen ions and hydrogen ions. In this
rocess a protonospheric field tube serves as a reservoir of ionization. There
s of course no barrier in the plane of the geomagnetic equator, and the ioni-
ation in the conjugate half of the field tube is part of the same reservoir.
xchange of ionization between the protonosphere and the F2-layer may also take
lace in the conjugate hemisphere, in this sense, the F2-layer in the conjugate
emisphere must also be regarded as a potential source or sink of ionization.

3.1.2 Aspects of F2-layer behavior to be Explained

The hypothesis will now be investigated that coupling between the ionosphere
nd the protonosphere is responsible for the well-known features of the F2-layer
he so-called seasonal anomaly in the critical frequency (that is, in the maxi-
um density of the layer), and the persistence of the layer throughout the night.

The seasonal anomaly in the critical frequency is one of the most puzzling
eatures of the F2-layer. It has been studied in detail by Yonezawa (1959 a,b)
nd by Croom et al. (1960). Under conditions of high sunspot number, the noon
ritical frequency in midsummer at a given location in medium or high geomagnetic
atitudes is found to be noticeably less than at midwinter in the same location.
hanges in temperature or composition of the high atmosphere have been suggested
Rishbeth, 1961) as possible explanations, but none have thus far been found
dequate.

Hanson and Ortenburger (1961) drew attention to the possibility of coupling
y charge exchange between the ionization in the protonosphere and the upper
art of the F2-layer. They concluded that this coupling would be weak at night

rthwell (1962) pointed out that the diffusion time of protons through atomic xygen ions, being impeded by coulomb forces, is highly temperature dependent, nd might therefore be expected to proceed much more rapidly during the day han at night. She therefore suggested that the decrease in summer hemisphere 2-layer electron density, and the increase in electron density in the opposite emisphere, might be due to diffusion of ionization along lines of the geomag- tic field between conjugate points in the two hemispheres. This explanation ill be stated here in the form of a hypothesis to be tested by the steady- tate theory of ionosphere-protonosphere coupling.

With the cessation of production of ionization by solar radiation at sunset, he F2-layer begins to decay. Two modes of decay are possible, depending upon he relative importance of diffusive transport. The location of the maximum ate of loss is a short distance below the layer maximum. If downward transport f ionization by diffusion is negligible, the layer decays from the bottom up- ard. This results in an increase of the altitude of the layer maximum with ime during the night. If, on the other hand, diffusive transport is efficient, onization diffuses downward to fill up the gap in the region of maximum rate f decay. This results in a layer which decays uniformly, the altitude of the aximum remaining relatively constant with time throughout the night. The ormer mode of decay seems to be most closely in agreement with observation, articularly during the post-midnight hours. However, the diffusion rate re- uired for this mode of decay is much slower than that expected from theore- ical considerations. In view of this discrepancy, it is necessary to consider he possible influence of phenomena other than diffusion and loss.

These phenomena may be divided into two categories: those which hinder he diffusion of ionization downward to a region of high loss rate, those which

replenish ionization. Insofar as they are responsible for an upward drift of ionization, atmospheric winds and electrodynamic forces belong in the first category. Replenishment of ionization may result from the existence of a flux of ionizing corpuscular radiation or from a source of ionization external to the layer.

The hypothesis to be adopted here is that the nighttime F2-layer is maintained by a flow of ionization into the top of the layer. The analysis of Hanson and Ortenburger (1961) shows that the coupling between the F2-layer and the protonosphere is weak at night, in the sense that only a small fraction of the ionization in a protonospheric field tube diffuses downward through the charge exchange region and into the F2-layer. However, the content of the protonosphere is sufficiently large that the possibility that the downward flow may be enough to maintain the nighttime F2-layer cannot be ruled out.

3.1.3 Purpose of the Chapter

The coupling between the F2-layer and the protonosphere is regulated by the rate at which hydrogen ions can diffuse through the region between the base of the protonosphere and the chemical equilibrium region for charge exchange with oxygen ions. The analysis of this process forms the subject matter of the next chapter. The purpose of this chapter is to investigate the effect of an external source or sink of ionization upon the structure of the F2-layer, and, in particular, to obtain an estimate of the magnitude of the influx or efflux of ionization which is required to produce a significant change in the structure of the layer.

The equations which govern the altitude dependence of the rate of production, loss and diffusive transport are presented in Section 3.2. The diffusion

equation relating the magnitude of the diffusive flow to the electron density distribution, is also derived there. When this relation is combined with the steady-state continuity equation, the latter takes the form of a second-order ordinary differential equation in the electron density. Solutions of this equation are presented, first, in Section 3.3, for a simple model in which the loss process is approximated by an ionization sink located below a fixed level, and second, in Section 3.4, for a distributed loss process. Electron density profiles corresponding to the solutions of Section 3.4 are presented in Section 3.5. The results of Section 3.4 are further employed in Section 3.6 to obtain a numerical estimate of the influx or efflux of ionization required to maintain or otherwise affect the electron density profiles deduced.

3.2 Formulation of the Theory

3.2.1 Altitude Variation of Physical Quantities

As previously pointed out, the neutral atmosphere is in a state of diffusive equilibrium above about 120 km. Thus, throughout the F-region, each constituent is distributed independently of the others according to the hydrostatic law

$$\frac{dp_r}{dz} = -n_r m_r g \tag{3.1}$$

where the subscript r is used to distinguish between the constituents. The neutral atmosphere at F-region heights is very nearly isothermal, and upon substitution for p_r in Equation (3.1) from the ideal gas law, there follows

$$\frac{dn_r}{dz} = -\frac{m_r g n_r}{kT} \tag{3.2}$$

The quantity

$$H_r = \frac{kT}{m_r g} , \tag{3.3}$$

is the scale height of the distribution of constituent r, and it is this quantity that is usually taken to define the altitude dependence of n_r. In the treatment to follow it is convenient to define the altitude dependence by the reciprocal of the scale height, denoted by a_r. Thus, upon integration of Equation 3.2, the concentration of constituent r is given by

$$n_r = n_{ro} \exp (-a_r z) \tag{3.4}$$

where the subscript zero denotes the concentration at the reference level $z = 0$.

The concentration of each of the constituents is not of direct concern except as it influences the altitude variation of the rate of production and rate of loss of ionization, and the rate at which this ionization can diffuse downward through the neutral atmosphere. The rate of production q has the altitude dependence of atomic oxygen, while the loss coefficient β has the altitude dependence of either molecular oxygen or molecular nitrogen. The rate at which diffusion proceeds for a given concentration gradient is governed by the diffusion coefficient D; the reciprocal of this quantity has the altitude dependence of the dominant neutral constituent present. The variation with altitude of these three quantities is defined by the constants a_1, a_2 and a_3:

$$q = q_o \exp (-a_1 z) \tag{3.5}$$

$$\beta = \beta_c \exp(-a_2 z), \tag{3.6}$$

$$D = D_c \exp(a_3 z), \tag{3.7}$$

A fourth constant a_4 defines the distribution that would be taken up by the electrons and accompanying oxygen ions under conditions of diffusive equilibrium:

$$n(O^+) = n_c(O^+) \exp(-a_4 z). \tag{3.8}$$

These four constants are defined exactly as in the paper by Bowhill (1962a).

It follows from Equation (2.19) that the mean ionic mass is in this case just one-half that of the mass of an oxygen atom. If this information is combined with the fact that atomic oxygen is the dominant neutral constituent in the F-region, it may be seen that, under normal circumstances,

$$a_1 = a_3 = 2a_4 = \text{constant}.$$

It is useful, however, to retain explicity all four constants in the analysis to follow.

3.2.2 The Diffusion Equation

The equation of motion of the F2-layer plasma is Equation (2.6), where the relative velocities are now those of the electrons and the atomic oxygen ions with respect to the neutral atmosphere. Because of the presence of the electric field, the electrons and ions diffuse together, a process known as ambipolar diffusion. Since the ion concentration is very much less than that of the neutral atmosphere in this region, the diffusion imparts no motion to the neutral atmosphere and it may therefore be regarded as stationary.

With these restrictions, along with the assumption of a plane-stratified atmosphere permeated by a uniform magnetic field in which diffusion is assumed to take place only along the lines of force of the field, the steady-state forms of the equations of motion become:

$$\frac{dp_i}{dz} - n_i (eE - m_i g) = \frac{kT_i n_i v_i}{D_{in} \sin^2 I} , \qquad (3.9)$$

$$\frac{dp_e}{dz} - n_e (eE - m_e g) = 0 . \qquad (3.10)$$

The subscript i has been used for the ions, and the quantity D_{in} is the diffusion coefficient for the diffusion of the ion gas through the neutral atmosphere. These equations may be compared with Equations (2.14) and (2.15) which govern the motion of a single ionic constituent through a mixture of ions and electrons.

Equations (3.9) and (3.10) may be further reduced by the same method used to reduce Equations (2.14) and (2.15) -- elimination of the electric field. This results in a single equation for the diffusion velocity v:

$$v = D_{in} \sin^2 I \, (1 + T_e/T_i) \left[\frac{1}{n_e} \frac{dn_e}{dz} + k \frac{m_i g}{(T_e + T_i)} \right] . \qquad (3.11)$$

This result is usually written in terms of the <u>ambipolar diffusion coefficient</u> D (hereafter referred to simply as the diffusion coefficient), defined by $D = 2D_{in}$. The second term in the square brackets in Equation (3.11) is simply

the reciprocal of the scale height for a mixture of oxygen ions and electrons when $T_e \neq T_i$. Equation (3.11) therefore takes the form

$$v = D \sin^2 I \frac{(1 + T_e/T_i)}{2} \left[\frac{1}{n_e} \frac{dn_e}{dz} + a_4 \right] . \qquad (3.12)$$

The sign convention here is such that v is positive for diffusion in the direction of decreasing z, that is, downward.

Two further simplifying assumptions will be made. One is that ion and electron temperatures are equal over the region of interest, an assumption common to all theoretical models of the F2-layer presented to date. The second assumption, adopted for convenience of notation, is that $\sin^2 I$ is unity; this condition is very nearly satisfied at middle geomagnetic latitudes.

3.2.3 The Continuity Equation

The electron density satisfies the continuity equation, which for a plane-stratified ionosphere takes the form

$$\frac{\partial n_e}{\partial t} = q - \beta n_e + \frac{\partial}{\partial z} \left\{ n_e (v + v_d) \right\} . \qquad (3.13)$$

Only steady-state solutions of this equation will be considered in this chapter. The quantity v_d is the drift velocity of the ionization imposed by electromagnetic forces of external origin and by motions of the neutral atmosphere; this will also be ignored.

The term flux is used in this and succeeding chapters to denote the rate of transport of particles or of heat across unit area. The downward flux G

of electrons due to diffusive transport is, from Equation (3.12),

$$G = D\,(n_e' + a_4 n_e) \qquad (3.14)$$

where differentiation with respect to z is denoted by a prime. As the flux term in Equation (3.13) is assumed to be due entirely to diffusive transport, substitution may be made for this quantity from Equation (3.14), with further substitution for q, β and D from Equations (3.5), (3.6) and (3.7), the steady-state continuity equation becomes

$$n_e'' - (a_3 - a_4)\,n_e' + a_3 a_4 n_e - (q_o/D_o)\exp\left[-(a_1 + a_3)z\right] - (\beta_o/D_o)\,n_e \exp\left[-(a_2 + a_3)z\right] = 0 . \qquad (3.15)$$

Solutions to this equation are considered in Sections 3.3 and 3.4.

3.2.4 Model Atmosphere and Boundary Conditions

The assumed model atmosphere enters the calculations only through the parameters a_1, a_2, a_3 and a_4. The two types of model atmosphere used in this chapter, referred to as types A and B, are defined in Table 3.1. The type A atmosphere is a special one used for illustrative purposes in Section 3.3. The type B model corresponds to an atmosphere in which the loss coefficient has the altitude dependence of molecular oxygen.

TYPE	a_1/a_4	a_2/a_4	a_3/a_4
A	2	$\gg 1$	2
B	2	4	2

Table 3.1 Relative magnitudes of parameters characterizing the model atmospheres of type A and type B.

The lower boundary condition which the solution of the continuity equation must satisfy is that $n_e \longrightarrow 0$ as $z \longrightarrow -\infty$. The upper boundary condition is a condition on the flux. In the absence of a source or sink of ionization above the layer, the upper boundary condition is that $G \longrightarrow 0$ as $z \longrightarrow \infty$. It follows from Equation (3.14) that this is equivalent to the condition that the distribution of ionization be proportional to $\exp(-a_4 z)$ at great heights. If, on the other hand, there is an influx or efflux of ionization $G(\infty)$ at the top of the layer, then the upper boundary condition is such that G, as given by Equation (3.14), tends to a finite value at great heights.

Solutions of the continuity equation satisfying the upper boundary condition of zero flux have been obtained for model atmospheres of both type A and type B by Bowhill (1962a). In this chapter solutions which satisfy the upper boundary condition of a non-zero flux will be investigated. This is a continuation of the work described by Bowhill (1962b).

The equilibrium distribution that prevails when all of the ionization is supplied by a downward flux of ions into the top of the layer will be called an accretive layer. The upper boundary condition for an accretive layer is that the flux tends to a limiting positive value. The same boundary condition applies to a layer in which ionization is provided by accretion and photoionization. The equilibrium distribution that exists under these conditions will be called a photoaccretive layer. Finally, the equilibrium layer that persists in a situation characterized by ionization flowing out at the top will be called a photodepletive layer. The upper boundary condition for this layer is that the flux tends to a limiting negative value.

3.3 Solutions for the Lossless F2-Layer

An idealized model of type A will be employed in this section. This model assumes a distribution of the molecular species such that all of the loss is essentially confined to a sink at and below the reference level $z = 0$. The lower boundary condition for the electron density is in this case $n_e = 0$, when $z = 0$. Solutions of the continuity equation for this case are presented only because their relative simplicity affords some physical insight into the process of layer formation. The accretive layer is discussed first, and it is shown that the maximum of this layer occurs at a fixed altitude independent of the magnitude of the incoming flux. This is followed by a discussion of the photoaccretive layer. Here another advantage accrues from use of the idealized model: a clearer presentation of the method of forming the solution of the continuity equation that satisfies the upper boundary condition for such a layer. Finally, discussion in terms of this model has the advantage of calling attention to the existence of a critical flux, the largest upward flux that can be supported by a diffusion process in an equilibrium layer.

3.3.1 The Accretive Layer

In the absence of production the continuity equation for the idealized model atmosphere takes the form

$$n_e'' - (a_3 + a_4) n_e' + a_3 a_4 n_e = 0 \qquad (3.16)$$

for all $z \geq 0$. The general solution of this equation is of the form

$$n_e = B \exp(-a_4 z) + C \exp(-a_3 z), \qquad (3.17)$$

where B and C are constants of integration. It is clear that the lower boundary condition $n_e = 0$ when $z = 0$, can be met only if $B = -C$. Further substitution for n_e from Equation (3.17) into Equation (3.14) provides an expression for the flux as a function of z

$$G = C\, D\, (a_4 - a_3) \exp(-a_3 z). \tag{3.18}$$

It follows from the assumed altitude dependence of the diffusion coefficient given by Equation (3.7) that the flux is at all altitudes constant and equal to $G(\infty)$, where, from Equation (3.18),

$$G(\infty) = -C\, D_o\, (a_3 - a_4) \tag{3.19}$$

The solution is therefore

$$n_e = \frac{G(\infty)}{D_o (a_3 - a_4)} \left[\exp(-a_4 z) - \exp(-a_3 z) \right] . \tag{3.20}$$

This distribution has the shape of layer with maximum density at a height

$$z_m = (a_3 - a_4)^{-1} \log (a_3/a_4) , \tag{3.21}$$

above the reference level, where the ionization sink is located. Evidently, this altitude is independent of the magnitude of the incoming flux, this is also a feature of the accretive layer for the model atmosphere with distributed loss, the type B atmosphere.

The distribution given by Equation (3.20) for the case $a_3 = 2a_4$ is shown in Figure 3.1. The layer distribution is the difference between two exponential

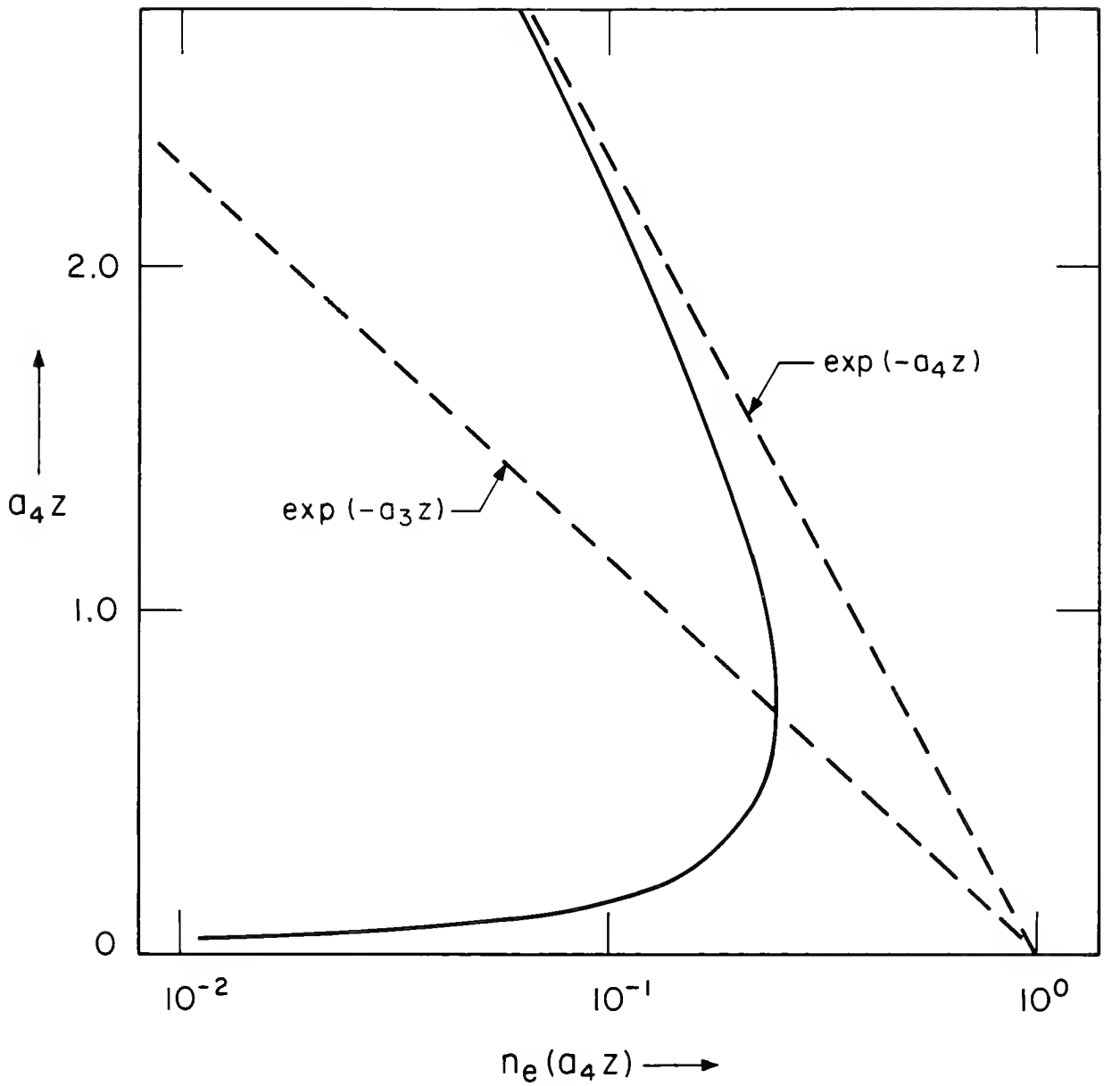

Figure 3.1 The continuous line represents the accretive layer distribution for the idealized model in which loss of ionization is confined to a sink at and below the reference level z = o. The electron density has been normalized such that the distribution at great heights tends to one of diffusive equilibrium with unit electron density at the reference level.

terms, each indicated here by a broken line. The term exp $(-a_4 z)$ has the altitude dependence characteristic of diffusive equilibrium and is just large enough to give $n_e \geq 0$ for all $z \geq 0$. The term exp $(-a_3 z)$ is required to sustain the downward flux. The flux is the same at all altitudes, but the distribution required to sustain this flux varies considerably with altitude. In the terminology of Chapter 2, the region above the layer maximum, where the distribution approaches one of diffusive equilibrium, is a region of fast diffusion. The distribution at and below the layer maximum departs considerably from that of diffusive equilibrium and is therefore a region of slow diffusion.

3.3.2 The Photoaccretive Layer

When the steady-state layer is maintained either entirely or in part by production of ionization within the layer, the continuity equation for the case of the idealized model atmosphere has the form

$$n_e'' + (a_3 + a_4)\, n_e' + a_3 a_4 n_e + (q_o/D_o) \exp\left[-(a_1 + a_3)\, z\right] = 0. \qquad (3.22)$$

The solution of this equation that satisfies both the lower boundary condition, $n_e = 0$, when $z = 0$, and the upper boundary condition that the flux tends to zero at great heights is

$$n_e = \frac{q_o}{D_o a_1 (a_1 + a_3 - a_4)} \left[\exp(-a_4 z) - \exp\left\{-(a_1 + a_3) z\right\}\right]. \qquad (3.23)$$

For the present purpose this function is to be regarded as a particular integral of Equation (3.22). The function defined by Equation (3.20) satisfies the homogeneous form of this differential equation and will therefore serve as the corresponding complementary function. A new general solution can be formed by adding

the particular integral and the complementary function to give

$$n_e = \frac{q_o}{D_c a_1 (a_1 + a_3 - a_4)} \left[\exp(-a_4 z) - \exp\left\{-(a_1 + a_3)z\right\} \right]$$

$$+ \frac{G(\infty)}{D_o (a_3 - a_4)} \left[\exp(-a_4 z) - \exp(-a_3 z) \right] . \qquad (3.24)$$

This solution of Equation (3.22) satisfies the upper boundary condition of a finite flux $G(\infty)$ and therefore gives the electron density as a function of altitude in a photoaccretive layer. Inspection of Equation (3.24) shows that, unlike the accretive layer, the shape of the photoaccretive layer depends upon the magnitude of $G(\infty)$. Elucidation of this dependence will be deferred until Section 3.5, where the distribution of electron density in the photoaccretive layer is calculated for the more physically realistic type B atmosphere.

Equation (3.24) also gives the distribution of electron density for the photoneutral layer. By definition, this layer is characterized by an efflux of ionization from the top of the layer, and $G(\infty)$ is therefore negative. Now at great heights Equation (3.24) gives

$$n_e \approx \left[\frac{q_o}{D_c a_1 (a_1 + a_3 - a_4)} + \frac{G(\infty)}{D_o (a_3 - a_4)} \right] \exp(-a_4 z) . \qquad (3.25)$$

A <u>critical flux</u> G_c may be defined by

$$G_c = - \frac{q_o (a_3 - a_4)}{a_1 (a_1 + a_3 - a_4)} \qquad (3.26)$$

It is evident that an upward flux of magnitude greater than G_c leads to a negative electron density at great heights. Physically this means that an upward flux of this magnitude cannot be supported by a diffusion process.

3.4 Solutions for the D2-Layer with Distributed Loss

In order to obtain the electron density distribution in a layer with loss processes distributed through it, the full equation

$$n_e'' - (a_3 - a_4)\, n_e' + a_3 a_4 n_e + (q_o/D_o) \exp\left[-(a_1 - a_3)z\right]$$

$$- (\beta_o/D_o)\, n_e \exp\left[-(a_2 - a_3)z\right] = 0 , \qquad (3.15)$$

must be solved. This equation is an inhomogeneous form of Bessel's differential equation. The substitutions which may be made to reduce the equation to the standard form, as well as the method of solution, are discussed by Bowhill (1962a) in greater detail than will be given here.

With the substitution for the independent variable z of

$$x = \exp\left[-(a_2 - a_3)z\right] , \qquad (3.27)$$

Equation (3.15) becomes

$$x^2 n_e'' + (1 - 2a)\, x n_e' + \left[\left(a^2 - \frac{\nu^2}{4}\right) - bx\right] n_e = x^\nu c \qquad (3.28)$$

where

$$\left.\begin{aligned} a &= \frac{a_3 + a_4}{2(a_2 + a_3)} \\ \gamma &= \frac{a_1 + a_3}{a_2 + a_3} \\ \nu &= \frac{a_3 - a_4}{a_2 + a_3} \end{aligned}\right\} \tag{3.29}$$

$$\left.\begin{aligned} b &= \frac{\beta_o}{D_o(a_2 + a_3)^2} \\ d &= \frac{q_o}{D_o(a_2 + a_3)^2} \end{aligned}\right\} \tag{3.30}$$

The further substitutions

$$\left.\begin{aligned} n_e &= Px^a \\ x &= -\frac{y^2}{4b} = \frac{y^2}{4b}\ \exp(\pi i) \end{aligned}\right\} \tag{3.31}$$

reduce the equation to the familiar Bessel form

$$y^2P'' + yP' + (y^2 - \nu^2)P = fy^{\mu+1}\ , \tag{3.32}$$

where

$$\left.\begin{aligned} f &= -4d(4b)^{a-\gamma} \exp\left[\pi i\ (\gamma - a)\right] \\ \mu &= 2\gamma - 2a - 1 \end{aligned}\right\} \tag{3.33}$$

Solutions of Equation (3.32) will be presented for the accretive layer, in Section 3.4.1, and for the photoaccretive layer, in Section 3.4.2. Electron density profiles calculated from these solutions will be discussed in Section 3.5.

3.4.1 The Accretive Layer

For the accretive case, $d = 0$, and Equation (3.32) then becomes

$$y^2P'' + yP' + (y^2 - \nu^2)P = 0 , \tag{3.34}$$

This is Bessel's differential equation; for non-integral values of ν the general solution is of the form

$$P = LJ_{-\nu}(y) + MJ_{\nu}(y) . \tag{3.35}$$

The lower boundary condition

$$n_e \rightarrow 0, \text{ as } z \rightarrow -\infty ,$$

transforms to

$$P \rightarrow 0, \text{ as } y \rightarrow \infty .$$

Inspection of the leading terms of the asymptotic expansion of $J_{-\nu}(y)$ and $J_{\nu}(y)$ shows that this condition requires

$$M = -L \exp(\nu\pi i) . \tag{3.36}$$

The expression for electron density then has the form

$$n_e = Lx^2 \left[J_{-\nu}(y) - \exp(\nu\pi i) J_{\nu}(y) \right] \tag{3.37}$$

As is well known, the Bessel functions have a representation as a power series

$$J_\nu(y) = \exp\left[\frac{-\nu\pi_1}{2}\right] \frac{(bx)^{\nu/2}}{\nu!} \sum_{m=0}^{\infty} \frac{\nu!}{m!(m+\nu)!} (bx)^m \quad , \tag{3.38}$$

$$J_{-\nu}(y) = \exp\left[\frac{\nu\pi_1}{2}\right] \frac{(bx)^{-\nu/2}}{(-\nu)!} \sum_{m=0}^{\infty} \frac{(-\nu)!}{m!(m-\nu)!} (bx)^m \quad . \tag{3.39}$$

For large z, small y,

$$x^a J_\nu(y) \approx \exp\left[\frac{-\nu\pi_1}{2}\right] \frac{b^{\nu/2}}{(\nu)!} x^{a+\nu/2} \quad a \exp(-a_3 z) \quad , \tag{3.40}$$

$$x^a J_{-\nu}(y) \approx \exp\left[\frac{\nu\pi_1}{2}\right] \frac{b^{\nu/2}}{(-\nu)!} x^{a-\nu/2} \quad a \exp(-a_4 z) \quad . \tag{3.41}$$

These forms may be substituted into Equation (3.37) to give an expression for the electron density that is valid at great heights. The upper boundary condition may be satisfied by substituting this limiting form for n_ϵ into Equation (3.14):

$$I(\infty) = D_c 1 \exp\left[\frac{\nu\pi_1}{2}\right] \left[\frac{a_3 b^{\nu/2}}{\nu!} - \frac{a_4 b^{\nu/2}}{\nu!}\right] \quad . \tag{3.42}$$

The upper boundary condition is satisfied if

$$1 = \frac{D(\infty)\nu! b^{-\nu/2}}{D_c (a_3 - a_4)} \exp\left[\frac{-\nu\pi_1}{2}\right] \tag{3.43}$$

The electron density in an accretive layer is therefore given by

$$n_e = \frac{G(\infty)\nu!b^{-\nu/2}x^{a}}{D_o(a_3-a_4)} \left[\exp\left[\frac{-\nu\pi i}{2}\right] J_{-\nu}(y) - \exp\left[\frac{\nu\pi i}{2}\right] J_{\nu}(y) \right] . \qquad (3.44)$$

In terms of the Bessel function of imaginary argument

$$I_{\nu}(s) = \exp\left[\frac{\nu\pi i}{2}\right] J_{\nu}\left(\exp\left[\frac{-\nu\pi i}{2}\right]\right), \qquad (3.45)$$

this may be written

$$n_e = \frac{G(\infty)\nu!b^{-\nu/2}x}{D_o(a_3-a_4)} \left[I_{\nu}(2\sqrt{bx}) - I_{-\nu}(2\sqrt{bx}) \right] . \qquad (3.46)$$

3.4.2 The Photoaccretive Layer

For the photoaccretive layer, the full equation

$$y^2P'' + yP' + (y^2 - \nu^2)P = fy^{\mu+1} , \qquad (3.32)$$

applies. The solution of this equation satisfying the upper boundary condition that the flux tends to zero at great heights has been obtained by Bowhill (1962a). The power series representation of this solution is given as Equation (A22) of that paper. There is a minor omission in that equation, though the correct version was used in the body of the paper, and the conclusions are unaffected. The correct equation reads:

$$n_e = \frac{dx^{\gamma}}{\lambda!(\lambda-\nu)!} \left[\frac{\lambda!(\lambda-\nu)!}{(-\nu)!(bx)^{\lambda}} \sum_{m=0}^{\infty} \frac{(-\nu)!}{m!(m-\nu)!} (bx)^m - \sum_{m=0}^{\infty} \frac{\lambda!(\lambda-\nu)!}{(\lambda+m)!(\lambda-\nu+m)!} (bx)^m \right], \qquad (3.47)$$

where

$$\lambda \equiv \frac{1}{2}(1+\mu+\nu) \quad . \tag{3.48}$$

As in Section 3.3.2 the solution satisfying the upper boundary condition for the photo-accretive layer is the sum of the accretive solution and the solution that satisfies the upper boundary condition that the flux tend to zero. Adding Equation (3.44) and (3.47), and making use of Equations (3.38) and (3.39):

$$n_e = \left[\frac{\lambda!(\lambda-\nu)!}{(-\nu)!\lambda(\lambda-\nu)}\, db^{-\lambda} + \frac{G(\infty)\nu! b^{-\nu}}{(-\nu)! D_o (a_3-a_4)}\right] x^{a-\nu/2} \sum_{m=0}^{\infty} \frac{(-\nu)!}{m!(m-\nu)!}(bx)^m$$

$$- \left[\frac{G(\infty)}{D_o(a_3-a_4)}\right] x^{a+\nu/2} \sum_{m=0}^{\infty} \frac{\nu!}{m!(m+\nu)!}(bx)^m$$

$$- \left[\frac{d}{\lambda(\lambda-\nu)}\right] x^{\gamma} \sum_{m=0}^{\infty} \frac{\lambda!(\lambda-\nu)!}{(\lambda+m)!(\lambda-\nu+m)!}(bx)^m \quad . \tag{3.49}$$

Direct comparison with Equation (3.24) is facilitated by noting that

$$\left.\begin{aligned} x^{a-\nu/2} &= \exp(-a_4 z) \\ x^{a+\nu/2} &= \exp(-a_3 z) \\ x^{\gamma} &= \exp -(a_1+a_3)z \end{aligned}\right\} \tag{3.50}$$

The critical flux is the value of $-G(\infty)$ that will make the first term in

brackets of Equation (3.49) equal to zero

$$G_c = \frac{\lambda!(\lambda-\nu)!}{\lambda(\lambda-\nu)} \frac{D_o(a_3-a_4)}{\nu!} db^{\nu-\lambda} \quad . \tag{3.51}$$

3.5 Calculation of Electron Density Profiles

Some examples of the solution of the continuity equation for the accretive layer and the photoaccretive layer will now be given. An atmosphere of type B will be assumed. Values of various non-dimensional quantities for the type B atmosphere that appear in Equations (3.44) and (3.47) are listed below in Table 3.2.

ν	γ	μ	λ
1/6	2/3	-1/6	1/2

Table 3.2. Values of non-dimensional quantities for the type B model atmosphere.

Upon substitution for the quantities listed in Table 3.2, the power series representation of Equation (3.44) becomes

$$n_e = \frac{G(\infty)}{D_o a_4} x^{1/3} \left[\frac{(1/6)!}{(-1/6)!} (bx)^{-1/6} \sum_{m=0}^{\infty} \frac{(-1/6)!}{m!(m-1/6)!} (bx)^m - \sum_{m=0}^{\infty} \frac{(1/6)!}{m!(m+1/6)!} (bx)^m \right] . \tag{3.52}$$

Expressing the first few terms of these series in numerical form one obtains

$$n_e = \frac{G(\infty)}{D_o a_4} x^{1/3} \left\{ \frac{(1/6)!}{(-1/6)!} (bx)^{-1/6} \left[1 + \frac{36bx}{5 \cdot 6} + \frac{(36bx)^2}{5 \cdot 6 \cdot 11 \cdot 12} + \cdots \right] \right.$$

$$\left. - \left[1 + \frac{36bx}{6 \cdot 7} + \frac{(36bx)^2}{6 \cdot 7 \cdot 12 \cdot 13} + \cdots \right] \right\} . \tag{3.53}$$

With the substitution

$$u = 36bx \ , \tag{3.54}$$

and substitution for b and d from Equations (3.30), the series (3.53) can be expressed in the form

$$n_e = G(\infty) \ (\beta_o D_o^2 a_4)^{-1/3} S(u) \ , \tag{3.55}$$

where

$$S(u) = \frac{(1/6)! 6^{1/3} u^{1/6}}{(-1/6)!} \left[1 + \frac{u}{5.6} + \cdots\right] - u^{1/3} \left[1 + \frac{u}{6.7} + \cdots\right] . \tag{3.56}$$

With the same substitutions, Equation (3.47) can be expressed in the form

$$n_e = \frac{q_o}{\beta_o} \left(\frac{\beta_o}{a_4^2 D_o}\right)^{1/3} T(u) \ , \tag{3.57}$$

where

$$T(u) = \frac{(1/2)!(1/3)! u^{1/6}}{(-1/6)!} \left[1 + \frac{u}{5.6} + \cdots\right] - u^{-1/3} \left[\frac{u}{2.3} + \frac{u^2}{2.3.8.9} + \cdots\right] . \tag{3.58}$$

The solution for the photoaccretive case is the sum of Equations (3.55) and (3.57):

$$n_e = \frac{q_o}{\beta_o} \left(\frac{\beta_o}{a_4^2 D_o}\right)^{1/3} T(u) + C(\infty) \ (\beta_o D_o^2 a_4)^{-1/3} S(u) \ . \tag{3.59}$$

With the aid of Equation (3.51) defining the critical flux, a common factor

can be extracted from the right hand side of Equation (3.59) to give

$$n_e = \frac{q_o}{\beta_o} \left(\frac{\beta_o}{a_4^2 D_o} \right)^{1/3} \left[I(u) + \frac{(1/2)!(1/3)!}{(1/6)! 6^{1/3}} \frac{G(\infty)}{G_c} S(u) \right] . \qquad (3.60)$$

The series S(u) and I(u) have been calculated for values of $\log_{10} u$ increasing from -4 to +3 by steps of 0.2.

Figure 3.2 shows the photoaccretive distribution for representative values of $G(\infty)$; the curve labelled $G(\infty) = - G_c$ is the photodepletive layer giving the largest outgoing flux that can be supported by a diffusion process. The abscissa is the electron density divided by the factor appearing outside the square brackets in Equation (3.60). Assuming $a_4 = (2H)^{-1}$, where H is the scale height of atomic oxygen, the ordinate of Figure 3.2 is related to the variable u by

$$\frac{z}{H} = - \frac{1}{3 \log_{10} e} \log_{10} u ,$$

the reference level having been chosen such that $z = 0$, when $u = 1$. The effect of a downward flux of ionization into the top of the layer is to increase the density at the layer maximum over that which is found in the absence of this flux. The dependence of the density at the layer maximum upon the magnitude of the flux is illustrated in Figure 3.3.

The electron density distribution for the accretive layer is given in Figure 3.4. The abscissa is the quantity S(u) appearing in Equation (3.55).

A frequently assumed approximate model for the shape of the F2-layer is the α - Chapman form (Wright 1960, Chandra 1963):

$$n_e = n_{eo} \exp \left[1/2 \left\{ 1 - z/H - \exp (- z/H) \right\} \right] . \qquad (3.61)$$

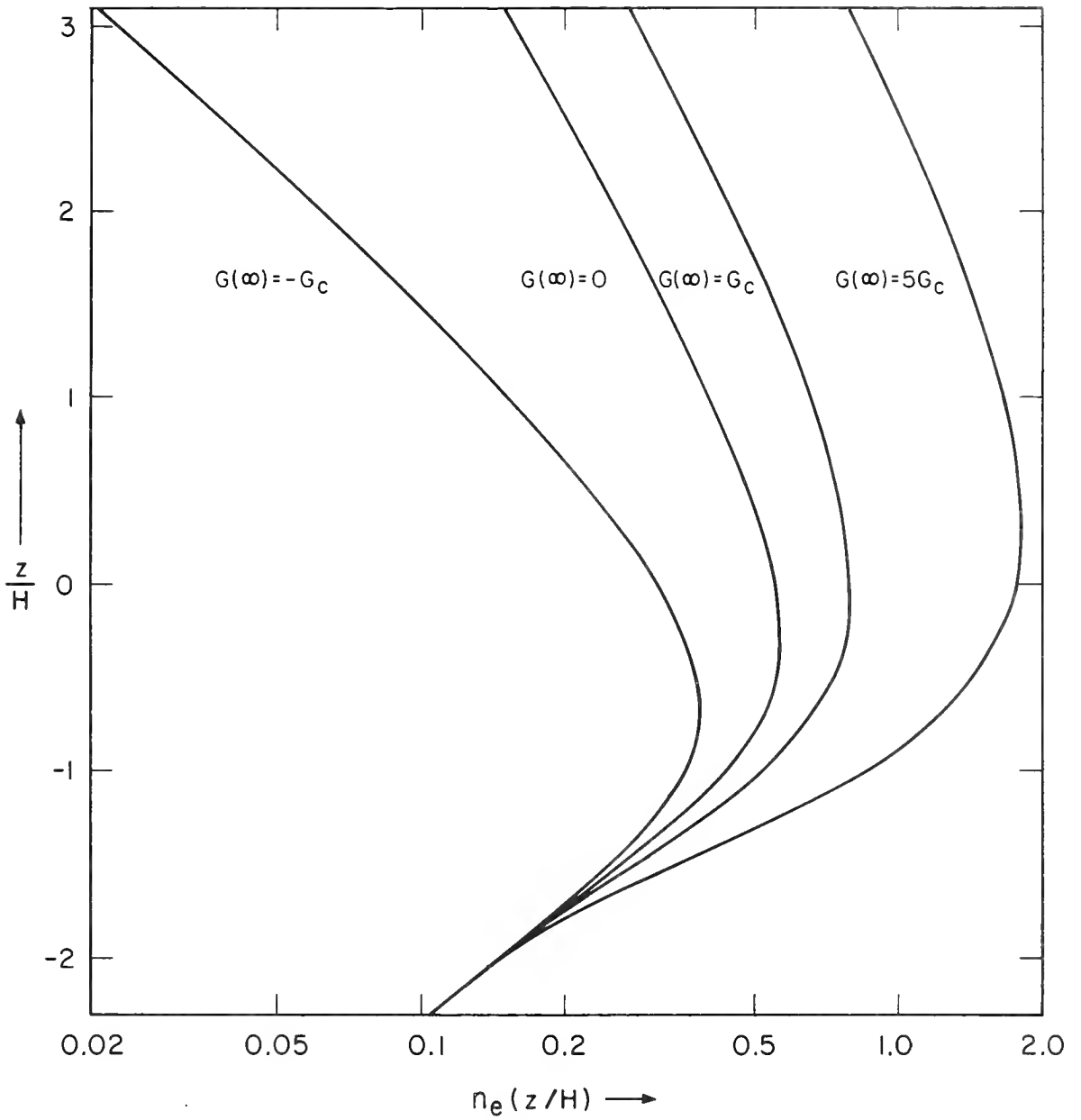

Figure 3.2 The photoaccretive layer distribution for various assumed values of the incoming flux (after Bowhill, 1962b). The curve labeled $G(\infty) = -G_c$ represents the photodepletive layer for the largest possible outgoing flux. The abscissa is electron density divided by the constant factor outside the square brackets in Equation (3.60).

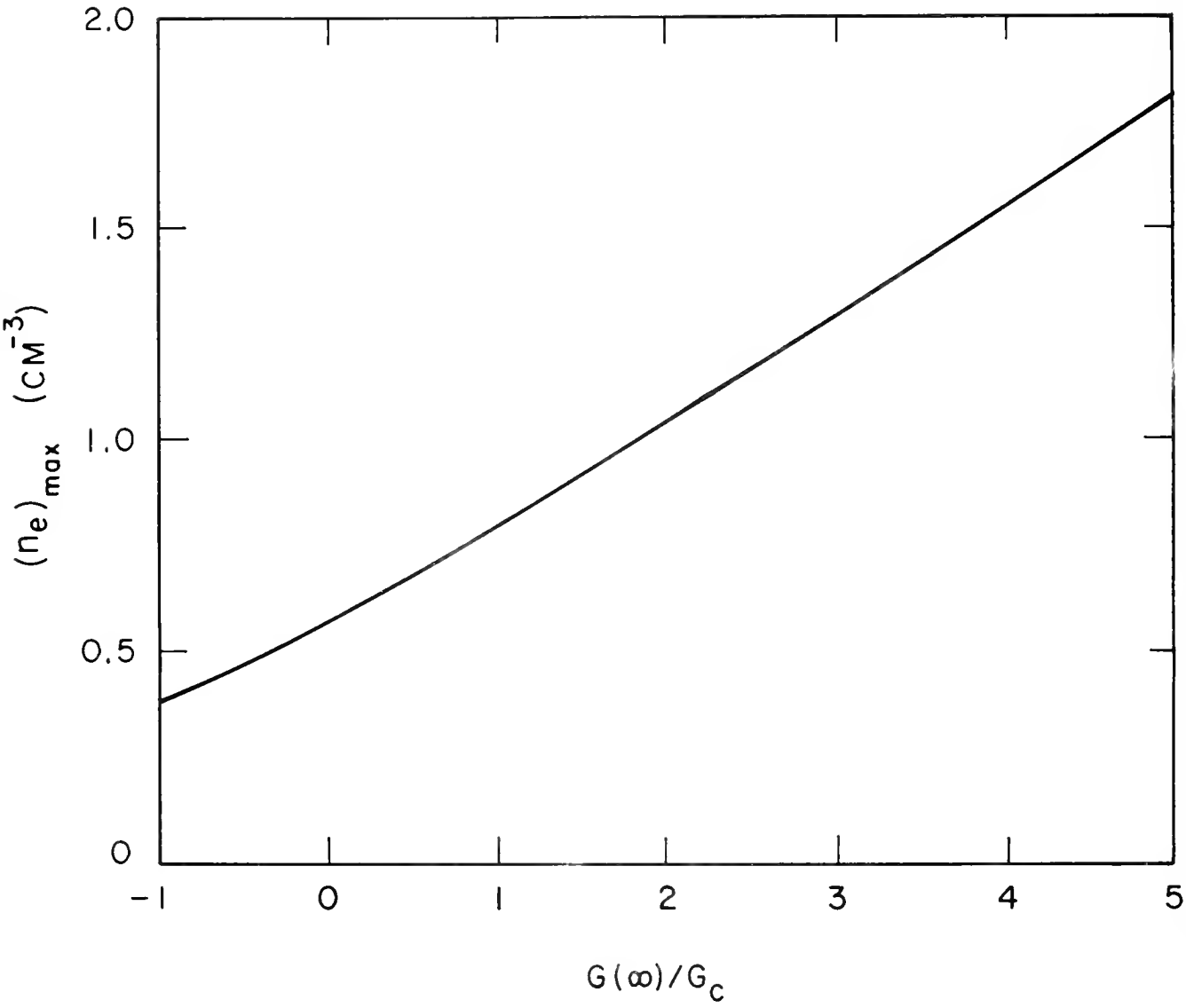

Figure 3.3 Dependence of the density at the maximum of the photoaccretive and photodepletive layers upon the magnitude of the flux at the top of the layer (after Bowhill, 1962b).

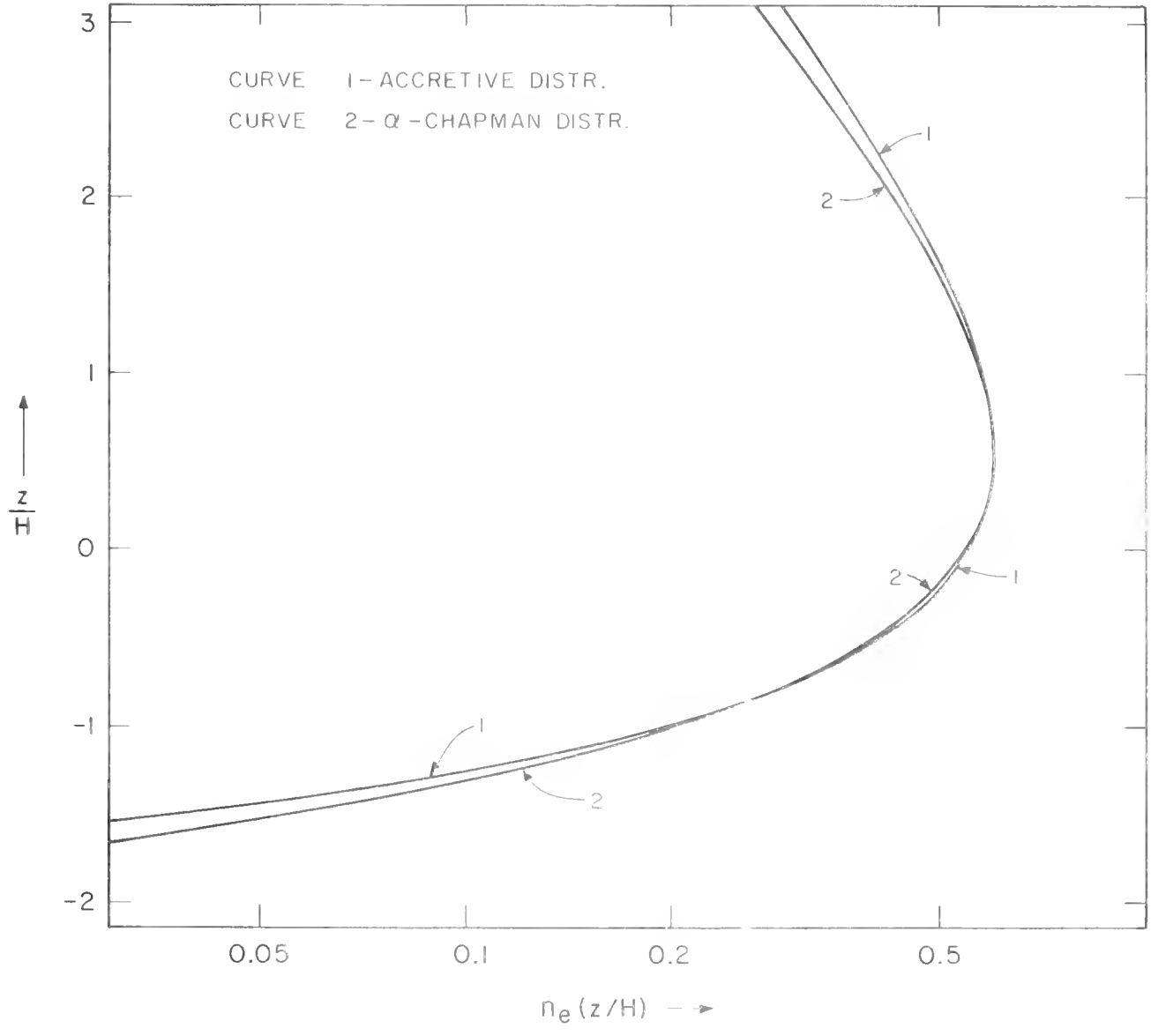

Figure 3.4 Comparison of the accretive layer distribution with the α-Chapman distribution having a maximum that coincides with that of the accretive layer (after Bowhill, 1962b). The abscissa is electron density divided by the quantity $G(\infty)\ (\beta_o D_o^2 a_4)^{-1/3}$ that appears in Equation (3.55).

This has an upper part corresponding to a diffusive equilibrium layer, and a rather sharp lower boundary. It is quite unlike the equilibrium F2-layer for the case $G(\infty) = 0$, which has an exponential tail on its lower side.

However, a comparison with the accretive layer (see Figure 3.4) indicates that this has a nearly α - Chapman form. The distribution of the α - Chapman layer having a maximum that coincides with that of the accretive layer appears in Figure 3.4.

3.6 Application of the Theory

3.6.1 The Daytime Layer

The hypothesis has been advanced that the seasonal anomaly is brought about through transfer of ionization from the summer hemisphere to the winter hemisphere by diffusion along the lines of the geomagnetic field. The analysis of this chapter shows that the rate at which ionization may be thus removed from unit column of the ionosphere cannot exceed that of the critical flux G_c. Inspection of Figure 3.3 shows that this process can increase the electron density at the F2-layer maximum in the winter hemisphere by at most a factor of 2 over that of the summer hemisphere. In order to account for the observed magnitude of the seasonal anomaly, the flux must be close to the critical value.

The critical flux is given by Equation (3.51). From the Equations (3.30) defining b and d and from Table 3.2, substitution may be made into Equation (3.51) to give

$$G_c = \left[\frac{(1/2)^{\frac{1}{3}}(1/3)^{\frac{1}{3}}}{(1/6)^{\frac{1}{3}} 6^{\frac{1}{3}} 3} \right] q_o \left(\frac{D_o}{\beta_c^2 4} \right)^{1/3} . \qquad (3.62)$$

For the type B model atmosphere, the factor $q_o (D_c/\beta_o a_4)^{1/3}$ is independent of altitude. With the assumption that $a_4 \simeq (2H)^{-1}$, where H is the scale height of atomic oxygen and after evaluation of the factor within the square brackets, the relation for the critical flux becomes

$$q_c = 0.468\, q \left(\frac{2HD}{\beta}\right)^{1/3} . \qquad (3.63)$$

The data to be used in the evaluation of the critical flux appear in Table 3.3. There are two values for q, β and D quoted for conditions representative of local noon near the minimum of the solar cycle, and near the maximum of the solar cycle. The values in parenthesis are all taken from Rishbeth (1964). The remaining values for q, β and D and for the other quantities have been taken from a variety of sources. All quantities listed apply to an altitude of 300 km.

The temperature and the concentration of atomic oxygen for sunspot minimum and sunspot maximum are taken from the Harris-Priester models S = 100 and S = 250, respectively. The scale height of atomic oxygen has been computed with this temperature from Equation (3.3). The quantity Dn(O) has been interpolated from the values quoted for several temperatures by Dalgarno (1964). The Dalgarno values are calculated for the diffusion of oxygen ions through atomic oxygen; these values have been doubled to obtain the ambipolar diffusion coefficient. Values of the loss coefficient are based on the values deduced by Nisbet and Quinn (1963) from observations of the decay of the nighttime F2-layer; the sunspot minimum value is that quoted by these authors for a temperature of $1000^{\circ}K$, while the sunspot maximum value is obtained by extrapolation

Quantity	Sunspot Minimum	Sunspot Maximum	Unit
T	1030	1875	^{o}K
n(0)	3.72×10^{8}	9.23×10^{8}	cm^{-3}
Dn(0)	1.42×10^{19}	1.86×10^{19}	$cm^{-1}\ sec^{-1}$
D	3.8×10^{10}	2.0×10^{10}	$cm^{2}\ sec^{-1}$
	(7.3×10^{10})	(2.0×10^{10})	
β	6.5×10^{-5}	6.0×10^{-4}	sec^{-1}
	(1.0×10^{-4})	(1.0×10^{-3})	
q	200	720	$cm^{-3}\ sec^{-1}$
	(100)	(850)	
H	60	110	km
G_c	1.8×10^{9}	2.9×10^{9}	$cm^{-2}\ sec^{-1}$
	(9.6×10^{8})	(3.0×10^{9})	

Table 3.3. Values of atmospheric quantities at time of local noon. Sources are given in the test. Last two entries in the table are calculated values of the critical flux.

on the assumption that the dependence of β upon the temperature is that inferred from the change in β between 700°K and 1300°K. The sunspot minimum value of q is roughly that deduced by Dalgarno et al. (1963) from the ionization cross section data and measurements of solar flux given by Watanabe and Hinteregger (1962). This value has been scaled by the corresponding ratio of the intensity of solar radiation at the 10.7 cm wavelength to obtain the sunspot maximum value.

Rishbeth (1964) has presented a time-dependent model of the F2-layer. The parameters of the model have been adjusted to give the best comparison with the observed behavior of the F2-layer. The values of q, β and D enclosed in parenthesis correspond to the values used in his "fast diffusion" model, which has been most successful in explaining the behavior of the daytime layer.

The values of G_c calculated from Equation (3.63) with these data are also shown in Table 3.3. On the basis of this result, it is reasonable to assign values to G_c of 1.5×10^9 cm^{-2} sec^{-1} at sunspot minimum and 3.0×10^9 cm^{-2} sec^{-1} at sunspot maximum.

3.6.2 The Nighttime Layer

The analysis of Hanson and Ortenburger (1961) has shown that a downward flux of oxygen ions into the top of the F2-layer should be expected at night. There is some evidence from the diurnal variation of whistler dispersion to substantiate this. The hypothesis made here is that this flux is sufficient to account for the persistence of the layer throughout the night.

The hypothesis may be stated a little more precisely if the detailed behavior of the nighttime layer is considered. The layer does in fact decay rather rapidly for a few hours after sunset, but as the night progresses this rate of

decay slows down. There is in fact evidence (Croom et al., 1959; Wright, 1962) that on a number of occasions the decay ceases entirely near local midnight, with electron densities remaining nearly constant thereafter. In the light of this evidence the hypothesis becomes, in the terminology of this chapter, a statement to the effect that the post-midnight F2-layer is an accretive layer.

Thus, the quantity of interest here is the magnitude of the flux $G(\infty)$ required to maintain an accretive layer with maximum density equal to the observed post-midnight value. This flux may be found from Equation (3.55). Again, for the type B model atmosphere, the parameters β and D appear in the expression in a combination that is independent of altitude, so that

$$G(\infty) = \frac{n_e}{S(u)} \left(\frac{\beta D^2}{2H} \right)^{1/3} . \qquad (3.64)$$

Inspection of Figure 3.4 shows that at the maximum of the layer, $S(u) \approx 0.6$. A figure of 1×10^5 cm^{-3} is consistent with the density at the layer maximum observed on the occasions characterized by the existence of a nearly constant post-midnight level (Rishbeth, 1964).

The data used to evaluate β and D are listed in Table 3.4. All values apply to time of local midnight and to an altitude of 300 km. Again, the values enclosed within the parentheses have been taken from Rishbeth (1964) and correspond to his fast diffusion model. The source for the value of each of the remaining quantities is the same as that for the corresponding quantities in Table 3.3.

The values of $G(\infty)$ calculated from Equation (3.64) with these data appear also in Table 3.4. The dependence upon level of solar activity is slight. The

Quantity	Sunspot Minimum	Sunspot Maximum	Unit
T	700	1430	^{o}K
n(O)	2.10×10^{8}	8.75×10^{8}	cm^{-3}
Dn(O)	4.14×10^{19}	1.66×10^{19}	$cm^{-1}\ sec^{-1}$
D	5.4×10^{10}	1.90×10^{10}	$cm^{2}\ sec^{-1}$
	(1.45×10^{11})	(2.2×10^{10})	
β	1.5×10^{-5}	2.0×10^{-4}	sec^{-1}
	3.0×10^{-5}	(6.0×10^{-4})	
H	40	83	km
$\Phi(\infty)$	3.0×10^{8}	2.4×10^{8}	$cm^{-2}\ sec^{-1}$
	(6.2×10^{8})	(4.3×10^{8})	

Table 3.4. Values of atmospheric quantities at time of local midnight. Sources are given in the text. Last two entries in the table are calculated values of the flux required to maintain the nighttime accretive layer.

mean of the two values calculated with the Bishbeth data is about 5×10^8 cm^{-2} sec^{-1}. The same value has been obtained by Bishbeth (1964) from analogue computer solutions of the continuity equation. The other two values calculated here are both about 3×10^8 cm^{-2} sec^{-1}. According to Bishbeth (1964), the diffusion coefficient for the fast diffusion model may be too large, but reduction by a factor larger than 3 would result in a discrepancy between the model and the features of the daytime layer. Inspection of Equation (3.64) shows that reduction of D by a factor of 3 reduces $G(\infty)$ by a factor of about 1.7. In view of this possibility the choice $G(\infty)$ 3×10^8 cm^{-2} sec^{-1} will be made for the downward flux needed to maintain the nighttime accretive layer.

4. DIFFUSIVE COUPLING BETWEEN THE IONOSPHERE AND PROTONOSPHERE

4.1 Introduction

The chemical reaction of charge exchange between atomic oxygen ions and atomic hydrogen is the source of the ionized hydrogen present in the topside. In Chapter 2 it was shown how the competing effects of gravitational separation and maintainence of charge neutrality account for the emergence of ionized hydrogen as the dominant constituent. It was pointed out there that the ionization within a protonospheric field tube can enter or leave only at the base of the tube, and that, as a consequence, the content of the field tube is determined by the concentration of hydrogen ions in the region of chemical equilibrium located below the base of the protonosphere on the field line which defines the axis of the tube. If the concentration in the region of chemical equilibrium is altered, the content of the field tube changes also. An estimate of the rate at which the content of the field tube can adjust to a given change in concentration in the region of chemical equilibrium is the primary objective of this chapter.

The assumptions adopted for the treatment of this chapter are stated at the outset, in Section 4.2. This is followed, in Section 4.3, by a qualitative discussion of the process which regulates the rate of transfer of ionization to and from the field tube: diffusion of hydrogen ions between the protonosphere and the region of chemical equilibrium. The distribution of hydrogen ions is, at all altitudes, governed by the continuity equation, the steady-state form of this equation for the hydrogen ions is derived in Section 4.4. The solution of this equation is employed to calculate the distribution of ionized hydrogen

for a number of different values of the assumed flux of ionization into or out of the protonosphere. The results show quantitatively the relation between the total content of a protonospheric field tube, the concentration in the region of chemical equilibrium, and the flux. The rate of diffusion in the upward direction is found to be limited to values less than a critical value. An estimate of the magnitude of this critical flux is obtained, and this is compared, in Section 4.5, with the fluxes needed to account for the seasonal anomaly and the maintenance of the nighttime F2-layer, respectively. Following this, in Section 4.6, a steady-state model for ionosphere-protonosphere coupling is set up for a 24-hour period to determine whether or not the magnitude of the average rate of exchange of ionization is close to that of the critical flux. The chapter is concluded with a discussion of factors which could be expected to significantly alter the estimated value of the critical flux.

4.2 Physical Assumptions

The results of this chapter are based on the solution of a form of the steady-state continuity equation for ionized hydrogen to be derived in Section 4.4. The assumptions which are needed to obtain this form of the continuity equation are listed below:

(i) The presence of helium ions in the topside will be ignored. The dependence of the results obtained upon this assumption is discussed in Section 4.7.

(ii) Hydrogen ions will be assumed to be a minor constituent. The derived form of the continuity equation is therefore valid only in the region where this condition exists. The method of extending the solution into the region where hydrogen ions become the dominant constituent is described in Section 4.4.

(iii) Ion and electron temperatures will be assumed equal to the temperature of the neutral atmosphere. This is further discussed in Section 4.7.

(iv) The drag force exerted on the O^+ ions by a diffusive flux of H^+ ions will be neglected. For fluxes of the order derived here, this force produces a negligible effect on the O^+ ion distribution in the region where these ions are the major constituent.

4.3 Qualitative Discussion of the Coupling Process

4.3.1 Description by Means of a Hypothetical Situation

The purpose of this section is to describe qualitatively the mechanism of exchange of ionization between the F2-layer and the protonosphere. Frequent reference will be made in Figure 4.1. The distribution adopted by each of the constituents in a mixture of ions of O^+ and H^+ in diffusive equilibrium is represented by a continuous line. The concentration of these constituents at the 500 km level and the temperature are the same as those for the diffusive equilibrium model of the topside in Chapter 2, representative of daytime conditions near sunspot minimum.

Consider now the situation which is created when the concentration of O^+ is decreased by a factor of 3 at all altitudes, to give the distribution represented by the broken line in Figure 4.1. If it is assumed that no change occurs in the relative abundance of neutral hydrogen and oxygen, the concentration of H^+ ions in the chemical equilibrium region is lowered by the same factor. There is now an excess of ionization in a protonospheric field tube and a downward flow of ionization through the base of the tube takes place. The H^+ ions diffuse downward through the O^+ ions into the region of chemical equilibrium. Accumulation of ionization would quickly make up the deficit

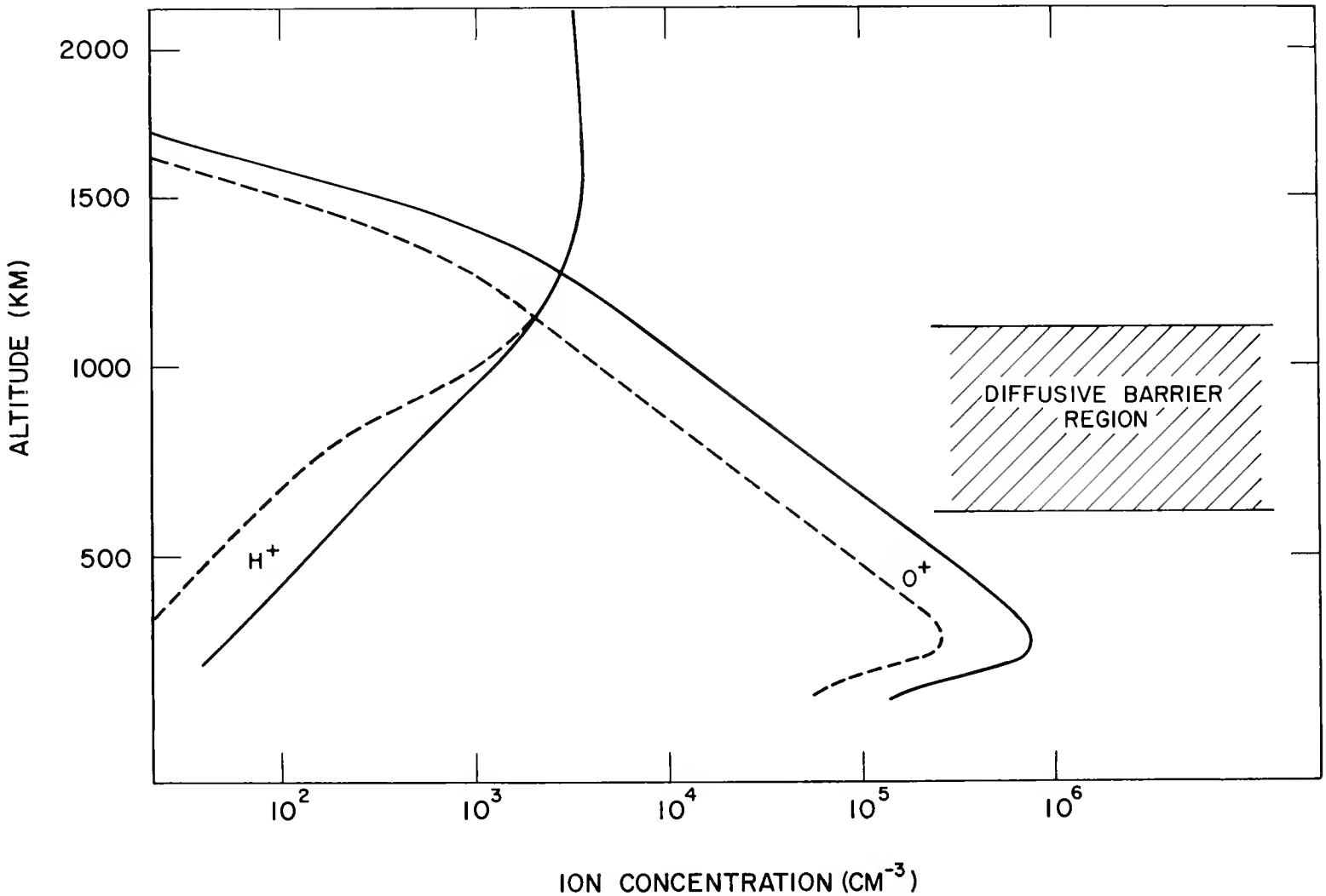

Figure 4.1 The continuous lines show the diffusive equilibrium distributions for ions of H^+ and O^+ when the concentration of these constituents at 500 km is as given in Table 2.1. The broken lines represent the distributions adopted when the concentration of O^+ is reduced by a factor of 3 at all altitudes. The H^+ ion distribution cannot immediately be restored to one of diffusive equilibrium because loss of ionization from the protonosphere is impeded by a diffusive barrier.

in this region and restore the distribution of ionized hydrogen to one of diffusive equilibrium were it not for the fact that the charge exchange reaction proceeds almost equally rapidly in either direction. Thus the H^+ ions which diffuse downward from the protosphere are converted to O^+ ions at a rate which is such as to maintain chemical equilibrium at altitudes below about 700 km. Another process by which the distribution of ionized hydrogen might be restored to one of diffusive equilibrium is the accumulation of O^+ ions sufficient to increase the concentration of these ions by a factor of 3 in the region of chemical equilibrium. This is not possible in the topside F2-layer however as O^+ ions are quickly removed by downward diffusion and subsequent recombination near or below the maximum of the layer.

It is therefore clear that a steady state situation will be set up very soon after the proposed decrease of the density of the F2-layer. This steady-state situation will be characterized by: (1) a downward diffusion of ionization through the base of the protospheric field tube, (2) conversion of H^+ ions to O^+ ions at a rate equal to the rate of supply from the field tube, and (3) a downward flux of O^+ ions into the F2-layer that is just equal to the flux through the base of the protospheric field tube. This situation is maintained until sufficient ionization is removed from the field tube to restore the distribution of ionized hydrogen to one of diffusive equilibrium at all altitudes above the region of chemical equilibrium.

The quantity of interest to be derived from the theory relating the processes described above is essentially the time required for restoration of a state of diffusive equilibrium to the ionized hydrogen. This depends upon the rate at which ionization can diffuse from the base of the protosphere down

to the level where the concentration is determined by chemical equilibrium. This region is one of slow diffusion and constitutes a diffusive barrier to the removal of ionization from a protonospheric field tube; the region is so marked in Figure 4.1. More precisely, the information needed to determine the degree of ionosphere-protonosphere coupling is the relation between the distribution of ionized hydrogen through the region of the diffusive barrier and the rate of change of protonospheric field tube content. In order to obtain this information, it is necessary to solve the continuity equation for ionized hydrogen.

The process of transfer of ionization from the protonosphere to the F2-layer described above may take place in reverse. Consider the situation which is created when the concentration of O^+ ions is at all altitudes increased by a factor of 3 over that of the distribution represented by the continuous line in Figure 4.1. There is now a deficiency of ionization within a protonospheric field tube. For the reasons stated above, a steady-state condition obtains in which ionized hydrogen flows upward through the diffusive barrier and into the field tube. This ionization is supplied at the expense of the content of the F2-layer. It was seen in Chapter 3 that there is a limit to the flux of ionization which can be supplied by the F2-layer (the critical flux), and it is therefore theoretically possible that the time required for the distribution of ionized hydrogen to be restored to one of diffusive equilibrium is in this case set by the F2-layer critical flux rather than by the rate of transport of hydrogen ions through the diffusive barrier. This question cannot be answered until full details have been obtained from the solution of the continuity equation for ionized hydrogen.

4.3.2 The Coupling Process and the Nighttime F2-Layer

In the foregoing example a variation of content of the F2-layer was assumed to be the cause of exchange of ionization between the ionosphere and the protonosphere. Now, the content of the F2-layer is in fact subject to strong solar control, a factor of 5 being a typical figure for the diurnal variation of the electron density at the layer maximum. As the electron density is lowest at night, a downward flux of ionization into the layer should be expected during this period. The question to be answered in this chapter is whether or not this downward flux is close to the figure derived in Chapter 3 for the flux required to maintain the assumed nighttime accretive layer.

The quantity of ionization which is thus removed from a protonospheric field tube during the night must be replenished by an upward flux of ionization during the day. This is a most important consideration, for it means that an estimate of the nighttime flux cannot be made from consideration of the nighttime situation alone. It is for this reason that a steady-state model for ionosphere-protonosphere coupling has been set up for a 24-hour period. This is described in Section 4.6.

4.3.3 The Coupling Process and the Seasonal Anomaly

If the seasonal anomaly is due to a diffusion of ionization along magnetic field lines, then it is necessary to suppose that in the summer hemisphere (where the low F2-layer densities are observed) there is a flux of ionization upward into the protonosphere and that in the winter hemisphere there is a flux downward into the F2-layer. Although this contradicts the pattern of normal daytime behavior just described, in which ionization diffuses upward in both hemispheres to replenish the protonosphere, it is not beyond the realm of possibility. Thus,

if the temperature of the protonosphere in the summer hemisphere is higher than that of the winter hemisphere, the entire tube may be filled by upward diffusion in the summer hemisphere to such an extent that it overflows into the conjugate F2-layer. In the steady state, continued transfer of ionization could then result from a departure of the distribution of ionization within the field tube from diffusive equilibrium maintained by the asymmetry of temperature in the two hemispheres, as suggested by Rothwell (1962).

It was pointed out in Chapter 3 that, in order for the seasonal anomaly to be thus accounted for, the magnitude of the rate of transfer of ionization must be close to that of the F2-layer critical flux. The hypothesis that a transfer of ionization is responsible for the seasonal anomaly may thus be tested by determining whether or not a flux of this magnitude can push through the diffusive barrier between the F2-layer and the protonosphere.

4.4 The Distribution of Ionized Hydrogen Through the Region of the Diffusive Barrier

4.4.1 The Diffusion Equation

The equations which govern the distribution of ionization in the topside ionosphere are Equations (2.14) and (2.15). If the assumption is made that He^+ ions are absent and that O^+ ions remain stationary in the presence of a diffusive flux of H^+ ions, these equations become

$$\frac{dp_1}{dz} - n_1 (eE - m_1 g) + \frac{kT}{D_{13}} \frac{n_1 n_3}{(n_1 + n_3)} v_3 = 0, \qquad (4.1)$$

$$\frac{dp_3}{dz} - n_3 (eE - m_3 g) - \frac{kT}{D_{13}} \frac{n_1 n_3}{(n_1 + n_3)} v_3 = 0, \qquad (4.2)$$

$$\frac{dp_e}{dz} + n_e (eE + m_e g) = 0. \qquad (4.3)$$

The subscripts 1 and 3 denote ions of O^+ and H^+, respectively. Use has been made of the fact that $D_{13} = D_{31}$; the $\sin^2 I$ term has been omitted. Equation (4.3) may be used as in Chapter 2 to express the electric field in terms of the mean ionic mass m_+. Upon further substitution for the partial pressures from the ideal gas law and neglect of the drag force exerted on the O^+ ions by the diffusing H^+ ions, Equations (4.1) and (4.2) become

$$\frac{dn_1}{dz} = n_1 \left\{ \frac{m_+}{2} - m_1 \right\} \frac{g}{kT} , \qquad (4.4)$$

$$\frac{dn_3}{dz} = n_3 \left\{ \frac{m_+}{2} - m_3 \right\} \frac{g}{kT} + \frac{n_1 n_3}{D_{13}(n_1 + n_3)} v_3 . \qquad (4.5)$$

Over most of the region of the diffusive barrier, H^+ ions are a minor constituent. For this case, $m_+ = m_1$ where m_1 is the mass of an atomic oxygen atom. It is convenient to work in terms of the atomic oxygen mass, thus expressing the altitude variation of constituents in diffusive equilibrium in terms of the scale height H of atomic oxygen. Under the assumption that H^+ ions are a minor constituent, Equation (4.4) can be integrated immediately to give the distribution of O^+ ions:

$$n_1 = (n_1)_o \exp(-z/2H), \qquad (4.6)$$

where the subscript zero denotes the concentration at the reference level $z = 0$. For the same assumption, Equation (4.5) gives for the diffusive flux $\bar{G}$ of H^+ ions:

$$\bar{G} = D_{13} \left\{ \frac{dn_3}{dz} - \frac{7}{16H} n_3 \right\} . \qquad (4.7)$$

It follows from Equation (2.9), which defines the diffusion coefficient, that $D_{13} \propto (n_1)^{-1}$ in the region where H^+ ions are a minor constituent, so that

$$D_{13} = (D_{13})_o \exp (z/2H) . \qquad (4.8)$$

4.4.2 The Equation of Continuity

The charge exchange reaction accounts for both the production and loss of ionized hydrogen. As pointed out in Chapter 2, the rate coefficient for the loss process is a fraction 8/9 of the production rate coefficient. Thus, the steady-state continuity equation for ionized hydrogen is

$$K \left\{ n_1 n(H) - \frac{8}{9} n_3 n(O) \right\} + \frac{d\bar{G}}{dz} = 0, \qquad (4.9)$$

where $n(O)$ and $n(H)$ are, respectively, the concentration of atomic oxygen and of atomic hydrogen. Substitution for the flux from Equation (4.7) gives for the continuity equation in the region where H^+ ions are a minor constituent:

$$n_3'' + \left\{ \frac{1}{2H} - \frac{7}{16H} \right\} n_3' + \left\{ \left(\frac{1}{2H} \right) \left(- \frac{7}{16H} \right) \right\} n_3$$

$$+ \frac{K}{(D_{13})_o} (n_1)_o \, n_o(H) \exp \left[-(1/2H + 9/16H)z \right]$$

$$- \frac{8}{9} \frac{K}{(D_{13})_o} n_o(0) \, n_3 \exp \left[-(1/H + 1/2H)z \right] = 0, \qquad (4.10)$$

where primes denote differentiation with respect to z. It has been assumed here that atomic oxygen and atomic hydrogen are in diffusive equilibrium, that is,

$$n(H) = n_o(H) \exp (-z/16H) , \qquad (4.11)$$

$$n(0) = n_o(0) \exp (-z/H) . \qquad (4.12)$$

Comparison of Equation (4.10) with the equation of continuity governing the distribution of ionization in the F2-layer, Equation (3.15), shows that the two equations are of the same form. If now the substitutions

$$\left.\begin{aligned} a_1 &= 9/16 \; H \\ a_2 &= 1/H \\ a_3 &= 1/2H \\ a_4 &= - \; 7/16H \end{aligned}\right\} \qquad (4.13)$$

are made for the parameters expressing altitude dependence, and the substitutions

$$\left.\begin{aligned} \bar{q}_o &= K\, n_o(H)(n_1)_o \\ \bar{\beta}_o &= \frac{8}{9}\, K\, n_o(0) \end{aligned}\right\} \qquad (4.14)$$

are made for the production rate and coefficient of the loss rate, Equation (4.10) becomes

$$n_3'' + (a_3 + a_4)\, n_3' + (a_3 a_4)\, n_3 + \frac{\bar{q}_o}{(D_{13})_o} \exp\left[-(a_1 + a_3)z\right] - \frac{\bar{\beta}_o}{(D_{13})_o}\, n_3 \exp\left[-(a_2 + a_3)z\right] = 0. \qquad (4.15)$$

The parameters a_1, a_2, a_3 and a_4 are defined exactly as in Chapter 3. Thus, it may be seen from Equations (4.14) that $\bar{q} \propto \exp(-a_1 z)$ and $\bar{\beta} \propto \exp(-a_2 z)$; Equation (4.8) shows that $D_{13} \propto \exp[a_3 z]$. Finally, when $\bar{G} = 0$, hydrogen ions are in diffusive equilibrium, and inspection of Equation (4.7) shows that $n_3 \propto [-a_4 z]$ for this case.

4.4.3 Solution of the Continuity Equation

Equation (4.15) will be solved subject to the upper boundary condition that the flux $\bar{G}$ of hydrogen ions approaches a finite limiting value of magnitude $\bar{G}_p$ as $z \longrightarrow \infty$. As both production and loss of hydrogen ions are negligible in the protonosphere, $\bar{G}_p$ is the flux through the base of the protonosphere, at the top of the diffusive barrier region. The flux $\bar{G}$ at any altitude is related to

the hydrogen ion distribution through Equation (4.7), where, as in Chapter 3, the sign convention is such that a downward flux is positive. The lower boundary condition is that the hydrogen ion distribution tend to one of chemical equilibrium. It is sufficient for this purpose to require that $n_3 \longrightarrow 0$ as $z \longrightarrow -\infty$.

These boundary conditions are the same as those imposed upon the solution of the continuity equation for the ionization in the F2-layer. Thus, the desired solution of Equation (4.15) is identical to that of the F2-layer continuity equation for the photoaccretive layer, Equation (3.49). This is not to say, however, that the distribution of ionized hydrogen is that of a photoaccretive layer, since the altitude dependence of the production rate, loss coefficient and diffusion coefficient all differ from that of the corresponding quantities for the F2-layer. The form which Equation (3.49) takes when substitution is made for the coefficients in accordance with Equations (4.13) and (4.14) is considered in Section 4.4.4.

A feature of Equation (3.49) is the existence of a value of the upward flux that is such that the right-hand side of the equation tends to zero at great heights. This was termed the critical flux, and the interpretation given there was that this is the largest upward flux that can be supported by a diffusion process in the F2-layer. As the distribution of ionized hydrogen is governed by the same equation, there is also a critical flux $\bar{G}_{pc}$ for diffusion of hydrogen ions. This is to be interpreted as the largest upward flux that can pass through the diffusive barrier between the region of chemical equilibrium and the base of the protonosphere. The critical flux $\bar{G}_{pc}$ is given by Equation (3.51), with appropriate substitution for the quantities on the right-

hand side of this equation. A numerical estimate of this quantity is obtained in Section 4.4.5.

4.4.4 Calculation of the Distributions

The distribution of ionized hydrogen through the region of the diffusive barrier will now be calculated for various assumed values of the flux $\bar{G}_p$ through the base of the protonosphere. Values of some of the nondimensional parameters of Equation (3.49) appropriate to the present situation are listed below in Table 4.1. These values follow from Equation (4.13).

ν	γ	μ	λ
5/8	17/24	3/8	1

Table 4.1 Values of nondimensional quantities in Equation (3.49) for the case of the hydrogen ion distribution.

Substitution of the values listed in Table 4.1 into Equation (3.49), determines the coefficients in the power series. As in Section 3.5, it is convenient to make a substitution for the independent variable:

$$\bar{u} = 64\, \bar{b}\, x, \tag{4.16}$$

where

$$\bar{b} = \frac{\bar{\beta}_o}{(D_{13})_o (a_2 + a_3)^2}, \tag{4.17}$$

$$x = \exp\left[-(a_2 + a_3) z\right]. \tag{4.18}$$

With the understanding that the dependent variable now represents the concentration of hydrogen ions, Equation (3.49) takes the form

$$n_3 = \frac{\bar{q}}{\bar{\beta}} \left\{ 1 \quad \frac{\bar{G}_p}{\bar{G}_{pc}} \quad R(\bar{u}) \right\}, \tag{4.19}$$

where

$$R(\bar{u}) = \frac{(5/8)!}{(-5/8)!} \; 2^{15/4} \left[1 + \frac{\bar{u}}{3.8} + \frac{\bar{u}^2}{3.8.11.16} + \cdots \right] -$$

$$- \bar{u}^{5/8} \left[1 + \frac{\bar{u}}{8.13} + \frac{\bar{u}^2}{8.13.16.21} + \cdots \right] . \tag{4.20}$$

It is instructive to express the concentration of hydrogen ions in terms of the concentration of the constituents which are active in the process of production and loss of these ions rather than in terms of the production and loss rates themselves. Substitution may be made for $\bar{q}$ and $\bar{\beta}$ from Equations (4.14) to put Equation (4.19) into the form

$$n_3 = \frac{9}{8} \frac{n(H)}{n(O)} \; n_1 \left\{ 1 + \frac{\bar{G}_p}{\bar{G}_{pc}} \quad R(\bar{u}) \right\}. \tag{4.21}$$

The diffusive equilibrium distribution follows from Equation (4.21), when the flux is set equal to zero:

$$n_3 = \frac{9}{8} \frac{n(H)}{n(O)} \; n_1 \tag{4.22}$$

Reference to Equation (2.4) shows that this distribution is identical to that which prevails in the region of chemical equilibrium.

The series given by Equation (4.20) has been calculated for values of $\log_{10} \bar{u}$ increasing from -3.0 to +3.0 by steps of 0.5. This quantity is related to the altitude z above the reference level through Equations (4.17) and (4.18). The reference level has been chosen to be the 400 km level. The assumed concentrations of atomic oxygen, atomic hydrogen, and atomic oxygen ions at this level are given below in Table 4.2.

$(n_1)_c$	$n_o(O)$	$n_o(H)$
3.8×10^5	1.5×10^8	3.0×10^4

Table 4.2. Assumed concentrations at the 400 km level.

The diffusion coefficient may be calculated from Equation (2.9). It is convenient for future reference to retain the temperature dependence explicitly. The calculation gives

$$D_{13} = \frac{(8.5 \times 10^7)}{n_1} T^{5/2} \qquad (4.23)$$

The expression for the rate coefficient K has been taken from Hanson and Patterson (1963) as

$$K = (3.10 \times 10^{-11}) T^{1/2} \qquad (4.24)$$

Equations (4.23) and (4.24) and the concentrations given in Table 4.2 have been used to obtain the distribution of hydrogen ions for several assumed values of the flux $\bar{G}_p$ from Equation (4.21), a temperature of 1250°K has been assumed. The resulting distributions are shown in Figure 4.2. The fluxes are expressed

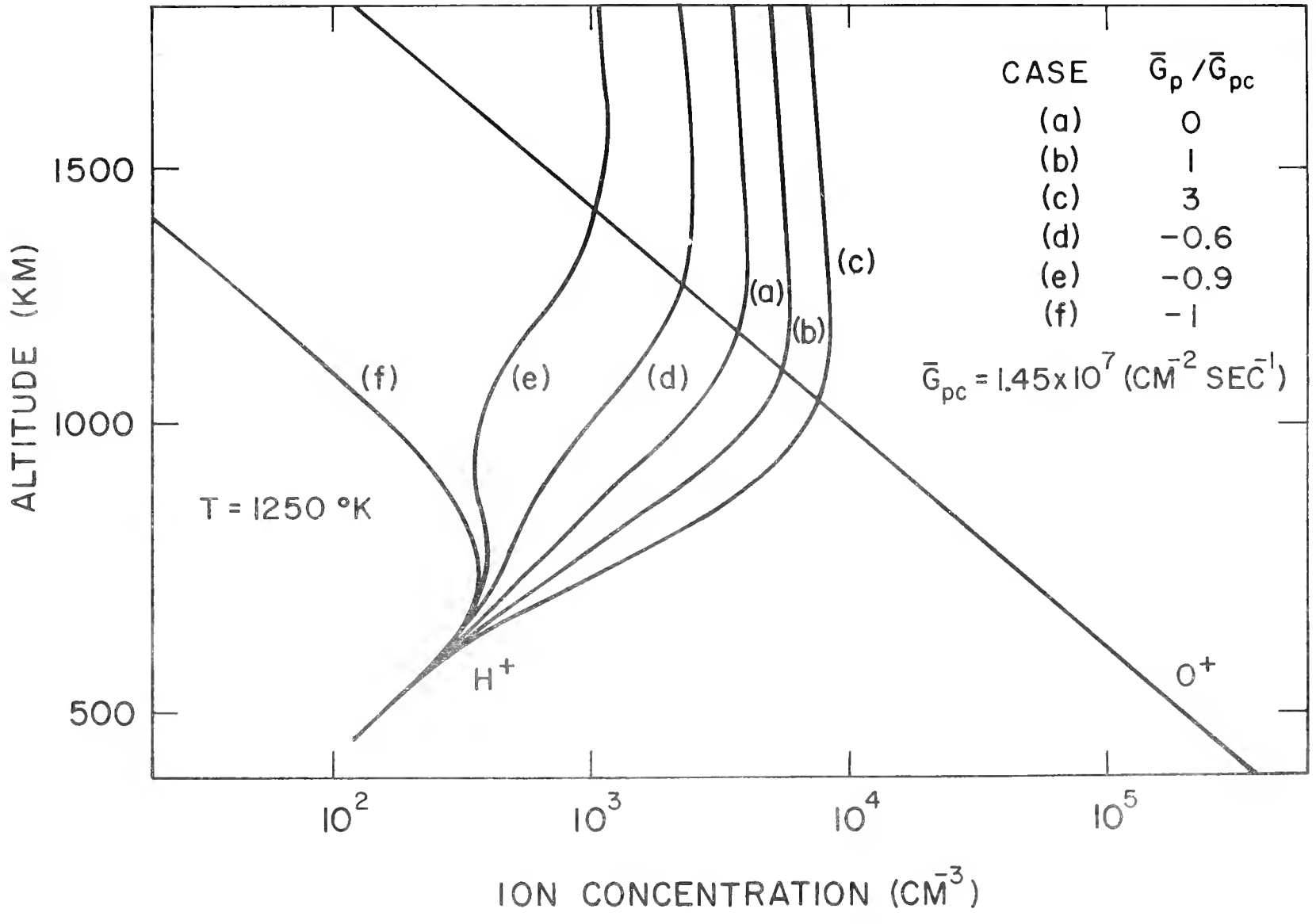

Figure 4.2 The distribution of ionized hydrogen which is required in each case to support the flux given in the upper right hand corner of the figure. Negative fluxes are upward. The concentration at 400 km is constant and specified by the quantities listed in Table 4.2.

in terms of the critical flux; a downward flux is positive. The curve (a) represents the diffusive equilbrium distribution and is the same as that of Figure 4.1. The remaining curves illustrate the relation between the distribution of ionized hydrogen through the region of the diffusive barrier and the rate of change of the content of a protonospheric field tube. As seen here, the rate of change of content depends on the difference between the concentration at the base of the protonosphere and the concentration at the top of the region of chemical equilibrium.

The distribution of O^+ ions in the region where these ions are a minor constituent is not shown in Figure 4.2; a different distribution prevails for each of the six cases shown here. The distribution that prevails when these ions are a major constituent has been assumed at all altitudes in order to show clearly the altitude of the level of transition to H^+ ions for each case.

Equation (4.21) was derived under the assumption that H^+ ions are a minor constituent. The distribution of these ions near and above the composition transition level therefore cannot be obtained from this equation. The method of continuing the solution up through the composition transition level and into the protonosphere is illustrated by Figure 4.3. The continuous line represents the solution of Equation (4.21) for case (e) of Figure 4.2. Above about 1200 km a straight line with slope equal to that of the diffusive equilibrium solution, case (a), is obtained. This indicates that, for an upward flux corresponding to that of case (e), the region above 1200 km is one of fast diffusion; that is, that the departure from diffusive equilibrium required to sustain this flux is small. The distribution of H^+ ions above 1200 km may

Figure 4.3 Illustrates the method of continuing the solution of the continuity equation for H^+ ions into the region where these ions are a major constituent. The solution of the continuity equation for H^+ ions is represented by the continuous line. The broken lines represent the diffusive equilibrium distribution of O^+ and H^+ ions above an assumed 1200 km reference level.

therefore be represented quite closely as one of diffusive equilibrium in a mixture of ions of O^+ and H^+. The distributions thus obtained for both O^+ and H^+ ions from Equation (2.25) with reference level at 1200 km are represented by the broken lines in Figure 4.3. With the exception of case (f), the solutions of Equation (4.21) all exhibit the behavior characteristic of a region of fast diffusion before the altitude of the composition transition level is reached. Thus, it is possible in each case to continue the solution upward through this level and into the protonosphere by the method illustrated in Figure 4.3. As mentioned previously, the distribution of O^+ ions obtained in each case is not shown in Figure 4.2.

4.4.5 The Critical Flux

The critical flux $\bar{G}_{pc}$ is given by Equation (3.51) with appropriate change in the notation:

$$\bar{G}_{pc} = \frac{\lambda!(\lambda-\nu)!}{\lambda(\lambda-\nu)} \, (D_{13})_o \, (a_3 - a_4) \, \bar{d} \, \bar{b}^{\nu-\lambda} , \tag{4.25}$$

where

$$\bar{d} = \frac{\bar{q}_o}{(D_{13})_o (a_2 + a_3)^2} , \tag{4.26}$$

and $\bar{b}$ is given by Equation (4.17). Upon substitution for $\bar{d}$ and $\bar{b}$ from these equations and for the remaining constants from Equations (4.13) and Table 4.1, Equation (4.25) becomes

$$\bar{G}_{pc} = \left[\frac{1}{(-3/8)!} \, \frac{(3/8)!}{(3/8)} \, (2/3)^{1/4} \right] \bar{q}_o \, H^{1/4} \left\{ \frac{(D_{13})_o}{\bar{\beta}_o} \right\}^{3/8} . \tag{4.27}$$

It is desirable to express the critical flux in terms of the atmospheric composition and the temperature. This may be accomplished by substitution from Equations (4.14) defining $\bar{q}_o$ and $\bar{\beta}_o$ and from Equations (4.23) and (4.24) for the diffusion coefficient and the rate coefficient to obtain

$$\bar{G}_{pc} = 3.85 \times 10^{-3} \left[\frac{n_o(H)(n_1)_o^{5/8}}{n_o(0)^{3/8}} \right] T^{3/2} \quad . \qquad (4.28)$$

Further substitution of the values given in Table 4.2 for the quantities within the square brackets of Equation (4.28) shows that $\bar{G}_{pc} = 1.45 \times 10^7$ cm^{-2} sec^{-1} for the example illustrated in Figure 4.2.

4.5 Discussion

4.5.1 The Seasonal Anomaly

If the seasonal anomaly is to be explained by diffusive transport of ionization between hemispheres, it is necessary to suppose that ionization is supplied to a protonospheric field tube in the summer hemisphere at a rate which is close to that of the F2-layer critical flux G_c. It was shown in Chapter 3 that this quantity is about 1.5×10^9 cm^{-2} sec^{-1} at sunspot minimum and higher by a factor of two at sunspot maximum. Factors which might alter significantly the value for the critical flux $\bar{G}_{pc}$ of 1.45×10^7 cm^{-2} sec^{-1} deduced above will be discussed in Section 4.7. It may be stated at this time, however, that present knowledge is such as to preclude the possibility of an increase in the calculated value of the critical flux $\bar{G}_{pc}$ by more than an order of magnitude over the value deduced above. The hypothesis that the seasonal anomaly is due to diffusive transport of ionization between hemispheres must therefore be rejected on the grounds that the critical flux $\bar{G}_{pc}$ for ionosphere-

protonosphere coupling is two orders of magnitude lower than the critical flux G_c for the F2-layer.

4.5.2 Maintenance of the Nighttime F2-Layer

It is evident from the discussion of Section 4.4 that there is no theoretical limit to the rate at which ionization may diffuse downward through the diffusive barrier between the protonosphere and the F2-layer. In view of the fact that the protonospheric field tube with foot at latitude 40^{o} contains some 4.5×10^{12} cm^{-2} hydrogen ions, it is possible that a downward flux sufficient to maintain the accretive layer model of the nighttime F2-layer presented in Chapter 3 could persist through most of the night. However, as pointed out in Section 4.3.2, the ionization thus removed from the field tube must be replenished by an upward flux during the day. As the upward flux cannot at any time during the day be larger than the critical flux $\overline{G}_{pc}$, it may be concluded that the magnitude of the nighttime flux cannot long be maintained at a value in excess of the daytime critical flux. As this quantity is an order of magnitude less than the value of 3×10^{8} deduced in Chapter 3 for maintenance of the nighttime accretive layer, it would appear that some other process must be invoked to explain the persistence of the nighttime layer.

However, as indicated in the preceding discussion of the seasonal anomaly, there is a possibility that a discrepancy of one order of magnitude can be reconciled by further analysis. In view of the fact that no other suggested mechanism for maintenance of the nighttime layer has been found to give better quantitative agreement with observation, such a step seems desirable. Before such an optimistic stand can be made, however, it is first necessary to

ascertain whether or not ionization is supplied to the field tube during the day at the maximum rate, that is, whether or not the magnitude of the daytime flux is close to that of the critical flux. For the purpose of obtaining an answer to this question, a steady-state model for the coupling process has been set up.

4.6 A Steady-State Model for Ionosphere-Protonosphere Coupling

4.6.1 Notation

In this section it will be necessary to make a distinction between daytime and nighttime conditions. In order to do this, further subscripting is necessary, and this makes it undesirable to continue usage of the convention wherein the concentrations of different ion species are distinguished by subscripts. In this section the symbols $n(O^+)$ and $n(H^+)$ will be used in place of n_1 and n_3, respectively to denote ion concentrations. The equations from which the steady-state model will be set up are summarized below for convenience of reference:

$$n(O^+) = n_o(O^+) \exp(-z/2H) , \tag{4.6}$$

$$n(H^+) = \frac{9}{8} \frac{n(H)}{n(0)} n(O^+) \left[1 + \frac{\bar{G}_p}{\bar{G}_{pc}} R(\bar{u}) \right] , \tag{4.21}$$

$$\bar{G}_{pc} = 3.85 \times 10^{-3} \left[\frac{n_o(H)\, n_o(O^+)^{5/8}}{n_o(0)^{3/8}} \right] T^{3/2} . \tag{4.28}$$

4.6.2 The Coupling Equation

The rate at which hydrogen ions flow through the diffusive barrier is determined by the difference between the concentration at the base of the protonosphere

and the concentration of these ions at a reference level chosen within the region of chemical equilibrium. A diurnal variation of the concentration at the base of the protonosphere results from the redistribution of ionization within a protonosphere field tube that accompanies a change of protonospheric temperature; this variation was calculated in Chapter 2. A diurnal variation of hydrogen ion concentration at a specified level within the region of chemical equilibrium results from the change of the ratio $n(H)/n(O)$ and the change of $n(O^+)$, at that level. The coupling equation, which is a relation between the H^+ ion concentration at the base of protonosphere and the O^+ ion concentration at the reference level in the region of chemical equilibrium, will be derived in this section. Following this, in Section 4.6.3, the coupling equations for the daytime and for the nighttime situation are combined into a single equation, using the condition that ionization taken from the field tube during the night must be replaced by day. This equation may be solved to give the ratio of the daytime flux to the critical flux.

The composition transition level is the level at which the concentrations of O^+ ions and H^+ ions are equal. This will hereafter be referred to simply as the transition level. For the treatment to follow, this level will be defined as the level at which $n(O^+) = n(H^+)$, when the distributions of these constituents are assumed to be given at all altitudes by Equations (4.6) and (4.21), respectively. Stated in another way, the transition level is here defined as the level at which the two continuous lines of Figure 4.3 intersect.

It was stated earlier that the altitude dependence of $n(H^+)$ given by Equation (4.21) tends to that of diffusive equilibrium as the transition level is approached from below. This is seen in Figure 4.3. The reason for this is

that $R(\bar{u})$ is very close to unity at these altitudes. Thus, if the subscript T is used to denote quantities evaluated at the transition level, there follows from Equation (4.21):

$$n_T(H^+) = \frac{9}{8} \frac{n_T(H)}{n_T(0)} n_T(0^+) \left[1 - \frac{\bar{G}_p}{\bar{G}_{pc}} \right] \qquad (4.29)$$

Upon substitution for the quantities on the right-hand side of this equation from the equations which define their respective altitude dependence one obtains

$$n_T(H^+) = \frac{9}{8} \frac{n_o(H)}{n_o(0)} n_o(0^-) \exp(7z_T/16H) \left[1 - \frac{\bar{G}_p}{\bar{G}_{pc}} \right] \qquad (4.30)$$

where the subscript zero denotes quantities evaluated at the reference level and z_T is the altitude of the transition level above this reference level. Now by definition of the transition level, $n_T(H^+) = n_T(0^-)$, and from Equation (4.6) the relation

$$\frac{n_T(H^-)}{n_o(0^-)} = \exp(-z_T/2H), \qquad (4.31)$$

may be obtained. From Equation (4.31) it follows that

$$\exp(7z_T/16H) = \left\{ \frac{n_o(0^-)}{n_T(H^-)} \right\}^{7/8} \qquad (4.32)$$

which may be used to eliminate z_T from Equation (4.30) to give

$$n_T(H^+)^{15/8} = K_c(\bar{G}_p + \bar{G}_{pc})\, n_o(O^+)^{10/8} \quad . \tag{4.33}$$

This is the coupling equation, which relates the H^+ ion concentration at the transition level to the O^+ ion concentration at the reference level. The quantity K_c is the coupling factor, given by

$$K_c = \left(\frac{27}{10}\right)\left(\frac{32}{81}\right)^{3/8} \frac{(5/8)!}{(-5/8)!} \left\{(D_{13})_o\, n_o(O^+)\right\}^{-3/8} \left\{K n_o(O)\right\}^{-5/8} H^{-1/4} \quad . \tag{4.34}$$

4.6.3 Specification of the Steady-State Model

The quantities appearing in the coupling equation are all functions of time, and a large number of solutions of this equation would be necessary in order to trace the magnitude and direction of the flux $\bar{G}_p$ over a 24-hour period. Such detailed treatment is not needed in order to provide the information sought here; namely, whether or not the difference between the average day and night values of $n_T(H^+)$ and of $n_o(H)/n_o(O)$ and $n_o(O^+)$ are such as to extract the maximum amount of ionization out of a protonospheric field tube that is permitted by the critical flux. The simplest assumption will be made: all physical quantities assume one value during the night and another during the day, the day and night being of equal duration. Under this assumption the condition that no change in field tube content occur over a 24-hour period imposes the condition that, if a flux $\bar{G}_p$ is present by night, then a flux $-\bar{G}_p$ is present by day. The possibility of coupling between hemispheres will be ignored.

For the simple steady-state model described above, only two coupling equations are needed, one for daytime conditions and one for nighttime conditions.

The quantity to be obtained from the solution of these equations is the ratio of the magnitude of the flux to the daytime critical flux denoted by $\bar{x}$. The daytime and nighttime coupling equations may be combined by division to form the single equation

$$\frac{(\bar{x} + \delta)}{(1 - \bar{x})} = \Omega \, , \tag{4.35}$$

where δ is the ratio of the nighttime value of the critical flux to the daytime value and

$$\Omega = \left[\frac{\left\{ n_T(H^+) \right\}_N}{\left\{ n_T(H^+) \right\}_D} \right]^{15/8} \frac{(K_c)_D}{(K_c)_N} \left[\frac{\left\{ n_o(O^+) \right\}_D}{\left\{ n_o(O^+) \right\}_N} \right]^{10/8} . \tag{4.36}$$

The subscripts D and N denote daytime and nighttime values, respectively. To avoid confusion, it should be pointed out here that the nighttime critical flux is not to be interpreted as a limit to the flux; it is simply the quantity obtained from Equation (4.28) when the nighttime values are substituted for the quantities on the right-hand side of this equation. Equation (4.35) may be solved for $\bar{x}$ to give

$$\bar{x} = \frac{1 - \delta/\Omega}{1 + 1/\Omega} \, . \tag{4.37}$$

4.6.4 Deduction from the Steady-State Model

The quantities appearing on the right-hand side of Equation (4.37) may be expressed in terms of composition and temperature if substitution is made for the diffusion coefficient from Equation (4.23) and for the rate coefficient

from Equation (4.24). These substitutions give the relations:

$$\frac{\delta}{\Omega} = \left[\frac{\left\{n_T(H^+)\right\}_D}{\left\{n_T(H^+)\right\}_N}\right]^{15/8} \frac{\left\{n_o(H)\right\}_N}{\left\{n_o(H)\right\}_D} \frac{\left\{n_o(O)\right\}_D}{\left\{n_o(O)\right\}_N} \left[\frac{\left\{n_o(O^+)\right\}_N}{\left\{n_o(O^+)\right\}_D}\right]^{15/8}, \quad (4.38)$$

$$\Omega^{-1} = \left[\frac{\left\{n_T(H^+)\right\}_D}{\left\{n_T(H^+)\right\}_N}\right]^{15/8} \left[\frac{\left\{n_o(O)\right\}_D}{\left\{n_o(O)\right\}_N}\right]^{5/8} \left[\frac{\left\{n_o(O^+)\right\}_N}{\left\{n_o(O^+)\right\}_D}\right]^{10/8} \left(\frac{T_D}{T_N}\right)^{3/2}. \quad (4.39)$$

A discussion of the behavior of the quantities δ/Ω and Ω^{-1} may best be introduced by consideration of the diffusive equilibrium situation. The relation between $n_T(H^+)$ and the concentration of O^+ ions and neutral constituents at the reference level that applies under conditions of diffusive equilibrium may be obtained from Equations (4.22) and (4.32):

$$n_T(H^+)^{15/8} = \frac{9}{8} \frac{n_o(H)}{n_o(O)} n_o(O^+)^{15/8}. \quad (4.40)$$

If substitution is made for $n_T(H^+)$ in Equation (4.38) from this relation, it may be seen by inspection that $\delta/\Omega = 1$; further reference to Equation (4.37) confirms that the flux is zero for this case. Thus, if the rate of exchange of ionization between the F2-layer and the protonosphere is to be very much less than the limit set by the critical flux, then the combination of day to night ratios given by Equation (4.38) must be close to unity.

The principal contribution to the right-hand side of Equation (4.38) comes from the oxygen ion term. The concentration of these ions at the F2-layer

maximum and in the region of chemical equilibrium is typically a factor of 5 greater in the daytime than at night, thus introducing a factor of about 1/25. The results of Chapter 2 indicate that the concentration of H^+ ions at the base of the protonosphere varies inversely as the temperature (see Figure 2.5), and, since the day-night ratio of this quantity is raised to the power 15/8 in Equation (4.38), another factor that is considerably less than unity is present. A recent study by Hanson and Patterson (1963) has shown that horizontal flow of atomic hydrogen is very efficient in preventing an appreciable day-night asymmetry in the abundance of this constituent; the atomic hydrogen term therefore has little effect in Equation (4.38). Only the atomic oxygen term introduces a factor greater than unity. However, the Harris-Priester models indicate that the day to night ratio of n(O) becomes appreciable only near the minimum of the solar cycle and even then is less than a factor of 5.

Thus, it is evident that δ/Ω is in fact much less than unity. This is due principally to the large diurnal variation of oxygen ion content. A further contribution is made by the fact that the diurnal variation of the concentration at the base of the protonosphere is phased with the variation of F2-layer density in such a way that the maximum possible difference between the hydrogen ion concentration at the base of the protonosphere and in the chemical equilibrium region is maintained over a 24-hour period.

The factor Ω^{-1} which appears in the denominator of Equation (4.37) would have to be very much larger than unity in order to have any effect on the flux. Inspection of Equation (4.39) shows that this is unlikely, since the concentration ratios of the ionic constituents again appear in a combination that produces a factor much less than unity. This is partially compensated for by

the temperature, but the day to night temperature ratio is not sufficiently large to alter the conclusion that Ω^{-1} is less than unity.

The quantities adopted for a calculation of the degree of ionosphere-protonosphere coupling based on the steady-state model are summarized in Table 4.3. The 400 km level has been chosen as the reference level. Quantities shown here are representative of conditions near the middle of the solar cycle. The temperature and concentration of atomic oxygen ions are taken from the Harris-Priester model S = 150 for local noon and local midnight. On the basis of the conclusion of Hanson and Patterson (1963) that the day-night asymmetry in atomic hydrogen abundance is small, the concentration of this constituent has been assumed constant and equal to the mean of the noon and midnight values tabulated in the Harris-Priester model used here. The figure of 1×10^4 cm^{-3} thus obtained is consistent with the value estimated by Hanson and Patterson (1963).

For the values appearing in Table 4.3, it is found that $\delta/\Omega = 0.043$ and $\Omega^{-1} = 0.12$. From the definition of $\bar{x}$ it then follows that $\bar{G}_p/\bar{G}_{pc} = 0.86$. It may therefore be concluded that the magnitude of the flux of ionization into the nighttime F2-layer is essentially equal to the upper limit imposed by the daytime critical flux.

The value of the critical flux for the model assumed here is 1.27×10^7 cm^{-2} sec^{-1}. This quantity is independent of any assumption about the content of a protonospheric field tube. From a consideration of the representative figure of 4.5×10^{12} cm^{-2} for the total content of a mid-latitude field tube derived in Chapter 2, it is evident that only a very small fraction of the ionization within the field tube participates in the diurnal exchange process.

This justifies the assumption that changes of concentration at the base of the field tube are due to temperature changes alone. The conclusion reached above that the rate of exchange of ionization is in fact very close to the critical flux $\bar{G}_{pc}$ is based on the ratio of day to night concentration at the base of the field tube thus deduced, but is independent of any assumption of the magnitude of the concentration. This quantity may in fact be deduced from the steady-state model presented here upon substitution for $\bar{G}_p$ into Equation (4.33). The values for day and night thus obtained are given in Table 4.4, along with $\bar{G}_p$ and the value of z_T obtained from Equation (4.32).

The values of $n_T(H^+)$ deduced from the model may be substituted for n_o in Equation (2.31) to deduce the concentration in the equatorial plane; the quantity r_o appearing in this equation is the geocentric distance to the transition level for each of the two cases. Values of 1.06×10^3 cm^{-3} and 9.2×10^2 cm^{-3} are thus found for the day and night cases, respectively, for an assumed geomagnetic latitude of 40^o; these are in very good agreement with the average value of 10^3 cm^{-3} deduced from whistler data for this latitude.

The ion distributions for this model are shown in Figure 4.4. The continuous lines represent the daytime distributions, while the broken lines represent the nighttime distributions. The departure of the distribution of H^+ ions from one of diffusive equilibrium is small at all altitudes during the night. This is not the case during the day; a rather large departure from the diffusive equilibrium distribution is found in the region of the diffusive barrier. This fact should be taken into account when measurements of H^+ ion concentration in this region are used to infer the altitude of the transition level or the concentration at some level in the region of chemical equilibrium.

Quantity		Day	Night
T	(^{o}K)	1500	1100
$n_o(H)$	(cm^{-3})	1.00×10^{4}	1.00×10^{4}
$n_o(O)$	(cm^{-3})	1.70×10^{8}	6.80×10^{7}
$n_o(O^+)$	(cm^{-3})	1.00×10^{6}	2.00×10^{5}

Table 4.3. Values of atmospheric quantities assumed for the steady-state coupling model.

Quantity		Day	Night
$n_T(H^+)$	(cm^{-3})	1.95×10^{3}	2.80×10^{3}
$\overline{G}_p$	($cm^{-2} sec^{-1}$)	-1.10×10^{7}	1.10×10^{7}
z_T	(km)	1650	1000

Table 4.4. Values of atmospheric quantities deduced from the steady-state coupling model.

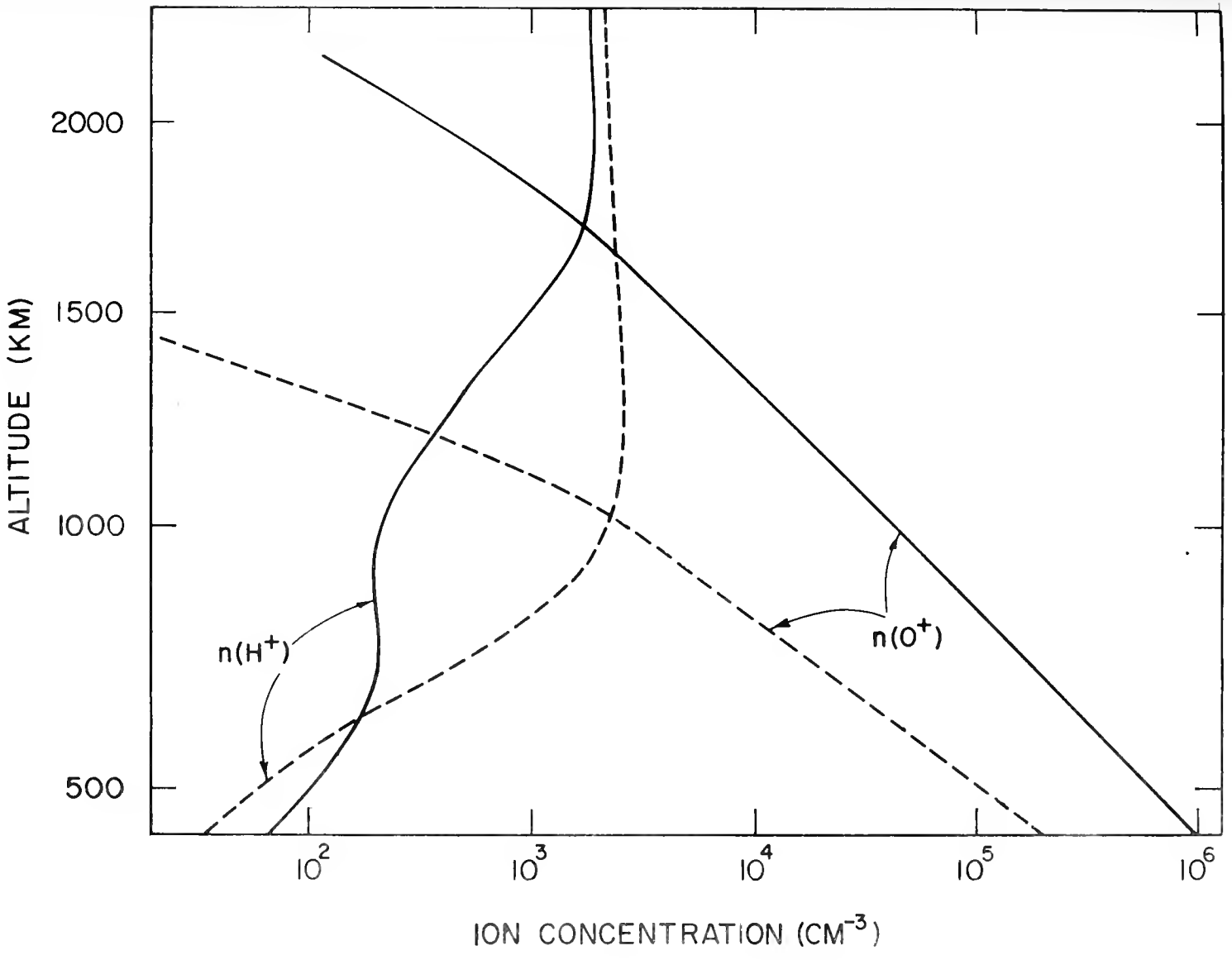

Figure 4.4 Ion distributions obtained from the steady-state model of ionosphere-protonosphere coupling for the values listed in Table 4.3. Continuous lines represent the daytime distributions, while broken lines represent the nighttime distributions.

4.7 Discussion of Factors which Influence the Critical Flux

It is convenient for the purpose of this section to write Equation (4.27), which defines the critical flux, in the form

$$\bar{G}_{pc} \propto \left\{ \frac{\kappa}{T^{1/2}} \right\}^{5/8} \left\{ \frac{D_{13}\, n(0^+)}{T^{5/2}} \right\}^{3/8} \left\{ \frac{n_o(H)}{n_o(0)^{3/8}} \right\} T^{3/2} \quad . \tag{4.41}$$

This form shows explicitly the dependence of the critical flux on the rate coefficient, the diffusion coefficient, the composition of the neutral atmosphere and the temperature. The first two terms in curly brackets are both constants, defined respectively by Equations (4.24) and (4.23). These quantities will be referred to subsequently as the rate coefficient constant and the diffusion coefficient constant. Frequent reference will be made to Equation (4.41) in the discussion to follow.

4.7.1 The Rate Coefficient

The value of 3.10×10^{-11} cm^3 sec^{-1} (^oK)$^{-1/2}$ adopted for the rate coefficient constant is taken from Hanson and Patterson (1963); according to these authors, this is probably an upper limit. A direct measurement of the distribution of H^+ ions by Taylor et al. (1963) has been used by Hanson et al. (1963) to deduce the rate coefficient at the time of the measurement. An atmospheric temperature of about 1250^oK was inferred from the slope of the oxygen ion distribution obtained from the same experiment. Division of the rate coefficient by the square root of the temperature gives a figure of 1.1×10^{-11} cm^3 sec^{-1} (^oK)$^{-1/2}$ for the rate coefficient constant. According to Hanson et al. (1963), this figure corresponds to a reaction cross section of 7.6×10^{-16} cm^2.

Calculations of this cross section based on laboratory measurements (Stebbings et al., 1964) indicate a value of about 1.0×10^{-15} cm^2 for a temperature of this order. The rate coefficient varies directly of the reaction cross section, and it would thus appear that the value deduced by Hanson et al. (1963) is a lower limit.

On the basis of the evidence given above it may be concluded that the rate coefficient constant is uncertain by at most a factor of 3. As this quantity enters the expression for the critical flux raised to the 5/8 power, an uncertainty of a factor of 2 is introduced into the derived value of the critical flux. This factor is small compared to that introduced by other quantities appearing in the expression for the critical flux.

4.7.2 The Diffusion Coefficient

The diffusion coefficient constant has been calculated from the expression for the diffusion coefficient derived by Chapman and Cowling (1958) for the process of binary diffusion of rigid spherical molecules with inverse-square repulsive forces. This diffusion coefficient has been assumed here to characterize the diffusion of H^+ ions through O^+ ions in the region where the former is a minor constituent. In view of the lack of theoretical analysis of this process, it is difficult to assess the correctness of this assumption. It may be mentioned that, in the calculation of the coefficient for the diffusion of O^+ ions through atomic oxygen (the diffusion process discussed in Chapter 3), explicit account must be taken of charge exchange (Dalgarno, 1958). This suggests that effects of charge exchange may contribute also to the process of diffusion of H^+ ions through O^+ ions. Discussion of the treatment of Chapman and Cowling is beyond the scope of this investigation and will not be attempted here.

4.7.3 Helium Ions

The diffusion process which couples the ionosphere and the protonosphere is in fact one of diffusion of H^+ ions through a mixture of ions of O^+ and of He^+. Although the theory described in this chapter is based on the assumption that He^+ ions are absent, it is possible to extend the theory to take into account the presence of these ions by a procedure which will now be briefly described.

It follows from Equation (2.9), which defines the diffusion coefficient D_{jk} that, when H^+ ions are a minor constituent,

$$\frac{D_{23}n_2}{D_{13}n_1} = \left(\frac{m_2 + m_3}{m_2 m_3}\right)^{1/2} \left(\frac{m_1 m_3}{m_1 + m_3}\right)^{1/2} , \qquad (4.42)$$

where subscripts 1, 2 and 3 denote ions of O^+, He^+ and H^+, respectively. Upon substitution for the mass ratios, the right-hand side of Equation (4.42) is found to be $(20/17)^{1/2}$, so that

$$D_{23}n_2 \approx D_{13}n_1 . \qquad (4.43)$$

The differential equation which governs the distribution of H^+ ions, Equation (4.5), may be modified to take into account the presence of He^+ ions by the addition of the He^+ ion drag term. The resultant equation is

$$\frac{dn_3}{dz} = n_3 \left\{\frac{m}{2} - m_3\right\} \frac{g}{kT} + \frac{n_1 n_3 v_3}{D_{13}(n_1 + n_3)} + \frac{n_2 n_3 v_3}{D_{23}(n_2 + n_3)} . \qquad (4.44)$$

It may be seen by inspection that this leads to an equation of the same form as that of Equation (4.7) which relates the flux $\bar{G}$ to the distribution of H^+ ions in the region where these ions are a minor constituent, provided D_{13} is replaced by an effective diffusion coefficient D_e defined by

$$\frac{1}{D_e} = \frac{1}{D_{13}} + \frac{1}{D_{23}} \quad . \tag{4.45}$$

With the aid of Equation (4.43), Equation (4.45) becomes

$$D_e = \frac{D_{13} n_1}{n_1 + n_2} \quad . \tag{4.46}$$

Reference to Figure 2.1 shows that at the 400 km reference level, $n_2 \ll n_1$, so that $(D_e)_o = (D_{13})_o$. Thus, the only effect of He^+ ions is to alter the altitude dependence of the diffusion coefficient.

To modify the theory of this chapter to allow for the presence of He^+ ions it is therefore necessary to change the value of a_3 from $1/2H$, which is the inverse of the scale height of the O^+ ion distribution, to a value which represents the average inverse scale height of the total ion distribution in the region where H^+ ions are a minor constituent. In view of the uncertainty in the relative abundance of He^+ ions present in the topside, this modification will not be attempted here.

The effect of the presence of He^+ ions on the critical flux may be seen from consideration of case (e) of Figure 4.2, for which the upward flux is close to the critical value. The region of the diffusive barrier here extends from about 700 km to about 1200 km. As may be seen from this figure, the

maximum departure from diffusive equilibrium occurs in the region below 1000 km. Consider now the effect of decreasing the diffusion coefficient by the same factor at all altitudes. In order to maintain the same flux, the distribution must depart still further from one of diffusive equilibrium, with the maximum change occurring below 1000 km. An equivalent statement is that for a given distribution, the change in the upward flux is governed primarily by the change in the diffusion coefficient below 1000 km.

For the model of the topside ionosphere shown in Figure 2.1, equal concentrations of O^+ ions and He^+ ions are found at about 1000 km. Reference to Equation (4.46) shows that the diffusion coefficient at the 1000 km level is thus decreased by a factor of 2 over the value for the case of no He^+ ions. However, due to the rapid decrease of He^+ ion concentration with decrease of altitude, this factor decreases to unity between 1000 km and 800 km. Reference to Equation (4.7) then shows that the presence of He^+ ions in the amount assumed here reduces the flux by at most a factor of 2.

4.7.4 Abundance of Atomic Constituents

The concentration of atomic hydrogen is probably the most uncertain quantity which appears in the expression for the critical flux. The reason for this is that atomic hydrogen is not present in the upper atmosphere in a sufficient amount for its relative abundance to be inferred reliably from satellite drag data. Such measurements do, however, provide sufficient information on the distribution of atomic oxygen that the effect of uncertainty in the concentration of this constituent on the critical flux is negligible.

The abundance of atomic hydrogen must be deduced from theoretical analysis of the following processes: (1) the rate of formation by dissociation of water

vapor near the 80 km level, (2) the rate of escape from the exosphere, and (3) the rate of diffusive transport between the 80 km level and the base of the exosphere. These problems have been considered by a number of investigators (Nicolet, 1957; Mange, 1961; Bates and Patterson, 1961; Kockarts and Nicolet, 1962). A fourth process, horizontal flow of atomic hydrogen between the day and night hemispheres, has been added in the recent analysis by Hanson and Patterson (1963).

The most uncertain element in the theory is the time-dependence of the rates listed above and, in particular, that of the rate of formation, which has been assumed independent of time of day and epoch of solar cycle in all treatments to date. At higher altitudes, the time-dependence is due to variation of temperature. Although the theory of escape has been adequately worked out, the resultant expressions for the escape rate are very sensitive functions of the temperature. The contribution of these effects to the day-night asymmetry of atomic hydrogen is moderated considerably by the conclusion of Hanson and Patterson (1963) that the rate of horizontal flow is an order of magnitude greater than the rate of supply from the lower atmosphere. The question of the dependence of hydrogen abundance on the epoch of solar cycle, however, remains unsolved.

4.7.5 The Ion Temperature

The assumption has been made in this chapter that ion and electron temperatures are equal to that of the neutral atmosphere. There is, however, considerable evidence of both a theoretical and observational nature that the ionized and the neutral components of the upper atmosphere are not in a state thermal equilibrium above about 120 km.

The thermal structure of the ionosphere has been investigated by Hanson and Johnson (1961), by Hanson (1963) and by Dalgarno et al. (1963). The results of these investigations indicate that the electron temperature at F-region heights during the day exceeds that of the neutral atmosphere by as much as a factor of 2, but that a state of thermal equilibrium is maintained between the ions and the neutral atmosphere.

The quantity T which appears in the equation defining the critical flux is the ion temperature in the region between 700 and 1000 km. To date no theoretical treatment of thermal equilibrium in this region has been carried out. Two interesting suggestions have, however, been put forward by Hanson (1963). The first of these is that, due to the low concentration of neutral constituents at these altitudes, ion and electron temperatures should be equal. The second suggestion is that a rather large fraction of the photo-electrons produced at F-region heights have sufficient energy to escape the ionosphere entirely, transferring much of their energy to the ionization in the protonosphere. It was further pointed out that the heat thus generated could be lost only through downward conduction by the electron gas to F-region heights.

The proposed downward conduction of heat by the electron gas requires that the electron temperature at the base of the protonosphere be higher than that at F-region heights. As the neutral atmosphere between the F-region and the protonosphere is isothermal, this would require a departure from thermal equilibrium at the base of the protonosphere. Dependent upon the rate at which the protonosphere is heated and upon the thermal conductivity of the electron gas, this departure may be very large. If so, the high ion temperatures

present would increase considerably the magnitude of the critical flux over that estimated in this chapter

It is for this reason that a full investigation of the thermal structure of both the ionosphere and the protonosphere has been undertaken. This research is described in the succeeding chapters.

5. THE THERMAL STRUCTURE OF THE IONOSPHERE

5.1 Introduction

The radiant energy absorbed in the thermosphere is of such short wavelength that much of it appears initially in the form of kinetic energy of photoelectrons. Some of this kinetic energy is transferred to the neutral constituents of the atmosphere through inelastic collisions and is responsible for the high temperature of the atmosphere that is observed in this region. A certain fraction of the initial photoelectron kinetic energy is, however, transferred directly to the ionospheric plasma through coulomb interactions. Thus, there exists a mechanism for selectively heating the ionized component of the atmosphere.

Because of the small electron-to-ion mass ratio, the amount of heat imparted to the ion gas in this manner is negligibly small compared to that imparted to the electron gas. The electron-to-neutral-particle mass ratio is also very small, and energy transfer between electrons is therefore much more efficient than energy transfer between electrons and ions or electrons and neutral particles. This makes it probable that the electron gas will be characterized by a kinetic temperature that is in general higher than that of the ion gas or the neutral atmosphere. Such a difference would be expected to be altitude-dependent as a consequence of the altitude dependence of the ratio of electron-to-neutral-particle concentration. In particular, the difference would be expected to be negligibly small in the lower ionosphere because of the very high electron-neutral particle collision frequency and consequent cooling efficiency.

The degree of this proposed lack of thermal equilibrium within the ionosphere has been the subject of several recent theoretical investigations (Hanson

and Johnson, 1961; Hanson, 1963; Dalgarno, McElroy and Moffett, 1963). The assumption common to all three of these investigations is that the rate of cooling of the electron gas is locally equal to the rate of heating at all altitudes. This provides an implicit equation for the electron temperature as a function of altitude. It should be noted that this assumption ignores the contribution of heat transport to the heat balance of the ionosphere.

Though differing somewhat in detail, the results of these investigations indicate a departure from thermal equilibrium beginning near 120 km, with the electron temperature increasing rapidly with altitude up to about 225 km. Here the electron temperature attains its maximum value, and this is roughly a factor of two greater than that of the neutral atmosphere. The electron temperature is found to decrease with further increase in altitude until thermal equilibrium is re-established near 350 km.

Since 1960 there have been a number of rocket-borne direct measurements of electron temperature within the ionosphere (Spencer, et al., 1962; Brace, et al., 1963; Nagy, et al., 1963). More recently, profiles of ion and electron temperature have been derived from incoherent backscatter measurements (Bowles, et al., 1962; Evans, 1962, 1964). Further direct measurements of electron and ion temperature have been obtained with the satellite Ariel I (Willmore, et al., 1964).

The rocket-borne direct measurements confirm the departure from thermal equilibrium near the 120 km level and show the rapid increase of electron temperature with altitude up to about 225 km. However, with rare exception, these measurements do not show the expected decrease of electron temperature above this altitude. Rather, a slow monotonic increase is indicated. This slow

increase is confirmed by mid-latitude backscatter profiles, and the results from Ariel I indicate that it continues to at least 1000 km.

It will be shown in this chapter that the discrepancy between the theoretically predicted profiles of electron temperature and the observed profiles is removed when explicit account is taken of the effects of heat transport by thermal conduction within the ionosphere.

The method of calculating the profiles of electron and ion temperature is presented in Section 5.7, and the results are discussed in Section 5.8. The intermediate sections are devoted to a quantitative discussion of the three competing physical processes that determine these temperature profiles: heating of the ambient electron gas by coulomb scattering of photoelectrons; cooling of the electron gas by collisions with the ions and with the neutral particles; and thermal conduction by the electron gas.

5.2 Local Heating

The rate of heat input to unit volume of the electron gas may be formally defined as

$$Q(z) = \int_0^\infty \epsilon(E,z)Ef(E,z)dE, \qquad (5.1)$$

where E is the initial energy of a photoelectron, ϵ is the fraction of this energy given to the electron gas, and $f(E,z)dE$ is the number of photoelectrons that appear per unit volume and unit time with initial energies in the range dE about E. A detailed discussion of the methods by which these functions are derived is given by Dalgarno, *et al.* (1963), and a brief description of some of the difficulties encountered is all that will be attempted here.

Calculation of the photoelectron distribution function f(E,z) demands some knowledge of the wavelength distribution of the solar ultraviolet photon flux and of the relevant photoionization cross sections. Only one measurement of the photon flux covering the region of interest (less than 1100 A) has been made to date (Damon, Hall, and Hinteregger, 1961). Quite apart from any uncertainty which may be inherent in an experiment of this type, there is the added disadvantage that the intensity and distribution of photon flux at some epoch of the solar cycle other than near the date of the measurement must be inferred by a method of questionable accuracy: scaling the flux over the entire spectrum by the ratio of the observed intensity at the 10.7 cm wavelength at the desired epoch to that near the time of the measurement. The principal difficulty in the analysis of the ionization processes is that the cross sectional data that are available for each of the absorbing constituents refer to the summed contribution from many levels, and it is necessary to make some assumption about the relative number of ionizing transitions that originate from a particular level of the participating constituent.

The loss of photoelectron energy through inelastic collisions with the neutral constituents involves many processes and few of the relevant cross sectional data are known. In practice, it is necessary to make some assumption concerning the relative number of excitational and ionizing collisions for photoelectron energies in excess of the ionization potential of a particular species. It is also necessary to assume some constant total cross section for the excitation of electronic states; the cross sections for excitation of vibrational and rotational states of N_2 and O_2 are better known due to the abundance of laboratory measurements.

Even when all the cross sectional data have been assembled, the determination of the fraction $\epsilon(E,z)$ for a photoelectron of initial energy E is no straightforward matter. This difficulty is inherent in the description of the behavior of a particle in terms of collision cross sections: it is very unlikely that two photoelectrons starting off with the same energy will make the same number of collisions of the same kind in an arbitrarily chosen segment of path. The nature of the difficulty can be seen more clearly if we consider for the moment an ensemble of systems, each consisting of unit volume of the atmosphere and ionosphere at altitude z, into each of which a photoelectron of energy E is introduced. In each system a different fraction of this energy will be given to the electron gas; the average of this quantity over the ensemble is the fraction $\epsilon(E,z)$.

The method of effecting such an average that has been adopted and carried out in varying amount of detail in all of the investigations cited can be outlined as follows: a model atmosphere is adopted, and the mean free paths for energy loss to the combined neutral constituents and to the electron gas are each plotted versus altitude. This is done for each initial energy, E. The altitude of the intersection of these two curves is then plotted versus E. This gives the critical energy as a function of altitude, in the sense that a photoelectron of this initial energy will, on the average, share its energy equally between the neutral atmosphere and the electron gas. The assumption is then made that at a given altitude the quantity $\epsilon(E,z)$ is unity for all photoelectrons ejected with energy less than the critical energy, and is equal to the ratio of critical energy to initial energy for all photoelectrons which are ejected with energy greater than the critical energy.

The functions $\epsilon(E,z)$ and $f(E,z)$ are then tabulated and the integration of Equation (5.1) yields an estimate of the rate of heat input to unit volume of the electron gas. Since the entire treatment is based tacitly upon the assumption that the mean free path of a photoelectron is much less than a length which is characteristic of the density variation in the medium through which it moves, the quantity defined by Equation (5.1) is referred to as the local heating rate.

5.3 Non-Local Heating

The treatment outlined in the preceding section is not entirely satisfactory due to the existence of a steep density gradient in the earth's atmosphere. It was first pointed out by Hanson (1963) that the mean free path for energy loss of a typical photoelectron exceeds the scale height of both the neutral and the ionized components of the atmosphere near the 300 km level. This has two implications. The first of these is that a certain percentage of the photoelectrons produced above this level will escape the ionosphere entirely, spiraling upward along magnetic field lines into the protonosphere. The second implication is that the region above the 300 km level is one of non-local heating; that is, a photoelectron deposits its energy over a distance that is large compared to the scale of density variation.

In attempting to formulate an expression for the heating rate in the region of non-local heating, it is necessary again to consider in some detail the mechanics of the energy loss processes for a single photoelectron. For this purpose we choose a "test" photoelectron and follow its path through the neutral and ionized components of the atmosphere. Effects of the magnetic field will be ignored for the time being.

The situation with regard to coulomb interactions is relatively straightforward. In the region of the ionosphere of interest here, the Debye length is much greater than the interparticle distance. The photoelectron therefore experiences a force field that is due to a large number of charged particles; close encounters are rare, and over a large distance the path of the photoelectron is nearly rectilinear. An expression for the rate of energy loss has been developed by Butler and Buckingham (1962). The initial energy of the vast majority of ionospheric photoelectrons is of the order of 10 electron volts, and for energies of this order it is permissible to use the asymptotic form of this expression (Dalgarno, et al., 1963):

$$\frac{dW}{ds} \sim \frac{-Kn_e}{W}, \qquad (5.2)$$

where $K = 1.95 \times 10^{-12}$ ev^2 cm^2, W is the energy of the photoelectron, n_e is the number density of the ambient electron gas, and s is the path length. The asymptotic form does in fact give satisfactory accuracy down to energies of one electron volt, and its use down to thermal energies does not lead to a very large error.

The simplicity of the situation just described stands in marked contrast to the complexity of energy loss to the neutral particles. In an elastic collision the photoelectron loses little energy, but the angle of scatter is usually large. In an inelastic collision the photoelectron may not only be scattered through a large angle, but also may lose a considerable fraction of its energy. Thus, the path of a photoelectron through the neutral atmosphere is very erratic, and an analytic expression for the energy loss per unit path length does not exist.

Now at altitudes above about 300 km the neutral gas density is sufficiently low that a large fraction of the initial energy of a photoelectron is delivered to the electron gas via coulomb interactions. It is therefore possible to develop with reasonable confidence a theory of non-local heating that ignores collisions between photoelectrons and neutral constituents. The effect of the neutral atmosphere may then be looked upon as a first order correction to the derived rate of heat input. This is the procedure that has been adopted here.

The model ionosphere is one in which the electron concentration $n_e(z)$ has a decreasing exponential dependence upon altitude with scale height 2H. Into this model ionosphere photoelectrons are ejected per unit volume with energy E at a rate q(z) that has a decreasing exponential dependence upon altitude with scale height H. The ionizable constituent is atomic oxygen; however, the presence of these neutral atoms will otherwise be ignored. The presence of atomic oxygen ions is understood insofar as they are necessary to preserve charge neutrality and thus account for the factor of two in the scale height of the electron density distribution.

A uniform magnetic field permeates the model ionosphere, the field lines making an angle I (the dip angle) with the surfaces of constant altitude. The assumption will be made that the pitch angle α does not change as the electron is slowed to thermal energy. It then follows from the equation of the spiral path and from the assumed validity of Equation (5.2) over the entire range of energy that the photoelectron loses energy according to the relation:

$$\frac{dW}{dz} = - \frac{K\, n_e}{W \sin I \cos \alpha} \qquad (5.3)$$

The geometry is illustrated by Figure 5.1.

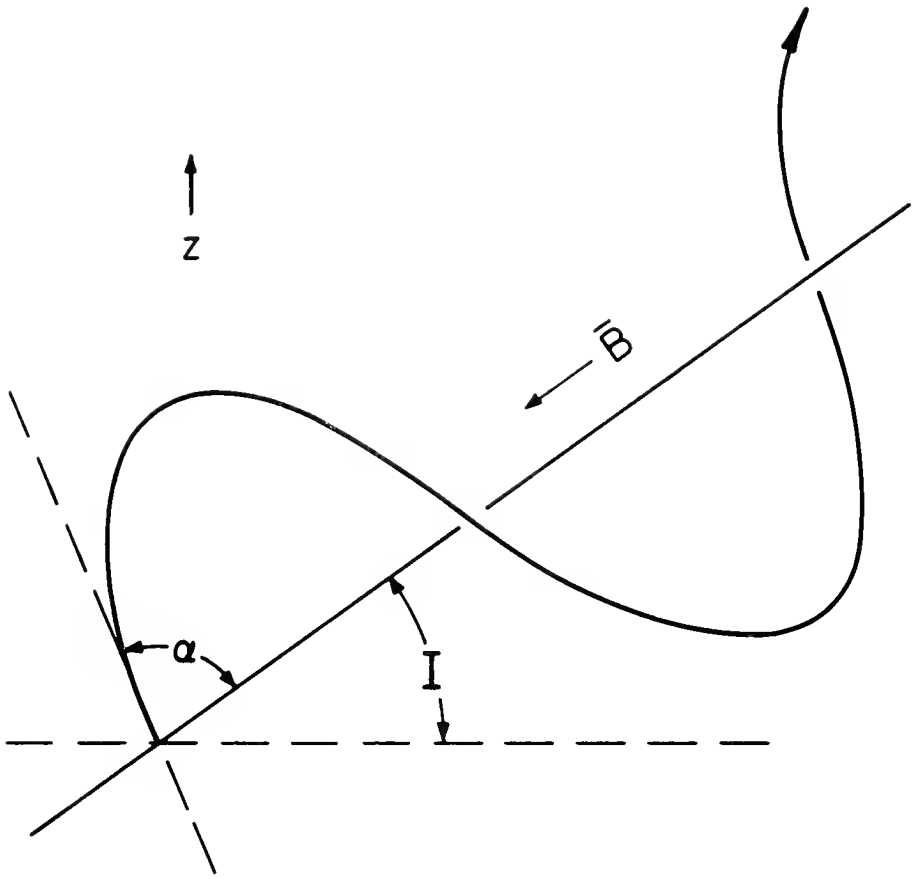

Figure 5.1 Path of a photoelectron with pitch angle α when the magnetic dip angle is I.

It is the purpose of Sections 5.3.1 and 5.3.2 to present in detail the derivation of expressions for the heating rate, the escape flux and the energy spectrum of the escaping photoelectrons. Two integrations are required in each expression: one over altitude and one over the pitch angle distribution. The differential expressions are rather complicated when the dependence upon pitch angle is included, and in the integration over pitch angle it is at times difficult to see what is taking place. For this reason, the theory is developed first in Section 5.3.1 without taking account of the effects of the pitch angle distribution. This will serve to clarify some of the physical concepts. In Section 5.3.2 the expressions are derived for an isotropic pitch angle distribution.

It should be pointed out here that the photoelectrons are not ejected isotropically. The correct initial pitch angle distribution is symmetric about a line in the direction of the sun (Mariani, 1964). The treatment appearing in Section 5.3.2 has been carried out for this pitch angle distribution also. It has been found that the resulting expressions show a rather weak dependence on solar zenith angle for moderate variations of this quantity, and give nearly the same results as the expressions derived for an isotropic distribution. As will subsequently be seen, the theory developed in this section is applied only to a noon situation, and the use of the isotropic distribution is therefore adequate. For situations of marked hemispherical asymmetry in the angle that the incoming radiation makes with the direction of the magnetic field (as, for example, at dawn near the summer solstice) the pitch angle distribution given by Mariani should be used.

5.3.1 Non-Local Heating by a Beam of Photoelectrons

In order to avoid for the present the complication introduced by the spiral motion of photoelectrons which are ejected with a velocity component perpendicular to the direction of the magnetic field, it will be assumed that of those photoelectrons produced in unit volume, one-half are ejected along the field line in the upward direction and the other half along the field line in the downward direction. The ambient electron gas in a particular volume is then heated by extracting energy from two beams of photoelectrons. The content of each beam is continually replenished by production of photoelectrons and depleted by thermalization. The number density of photoelectrons is at any point very much less than the number density of ambient electrons, and interaction between the two beams is negligible.

A short digression may conveniently be inserted at this point for the purpose of introducing two basic concepts: the photoelectron range and the escape level.

Consider a photoelectron which originates at level z with energy E and moves along a field line in the upward direction. If it is assumed that the electron concentration is independent of altitude, then it follows from integration of Equation (5.3) that the energy of the photoelectron at level z' is given by

$$W = E\left[1 - \frac{(z' - z)}{R}\right]^{1/2},$$

where

$$R = \frac{E^2 \sin I}{2Kn} . \qquad (5.4)$$

The quantity R is termed the range of the photoelectron, since when $(z' - z) = R$, $W = 0$. If now the altitude dependence of the electron concentration is included, integration of Equation (5.3) results in the relation

$$W = E\left[1 - \frac{2H}{R(z)}\left\{1 - \exp\left[-(z' - z)/2H\right]\right\}\right]^{1/2}, \qquad (5.5)$$

where R(z) is the range in a uniform ionosphere of density equal to that at the level from which the photoelectron originated.

If the photoelectron has sufficient initial energy, it may escape the ionosphere entirely. In such case z' becomes infinite, the exponential term in Equation (5.5) vanishes, and the energy carried away at escape is given by

$$W = E\left[1 - \frac{2H}{R(z)}\right]^{1/2}. \qquad (5.6)$$

If this energy is precisely zero, then the level from which the photoelectron originated is called the escape level, denoted by z_e. It follows from Equation (5.6) that $R(z_e) = 2H$, and this provides a relation between the initial energy and the number density at the escape level:

$$n(z_e) = \frac{E^2 \sin I}{4HK}. \qquad (5.7)$$

By definition, the rate of heat input $Q(z')$ to unit volume of the ambient electron gas at a level z' is the product of the total photoelectron flux through this level and the quantity $\left|\frac{dW}{dz}\right|$, integrated over the altitude range in which photoelectrons reaching level z' originate:

$$Q(z') = \left[\int_{z_1}^{z'} \frac{q(z)}{2} \left| \frac{dW}{dz} \right|_{z'} dz + \int_{z'}^{z_2} \frac{q(z)}{2} \left| \frac{dW}{dz} \right|_{z'} dz \right] . \qquad (5.8)$$

Some discussion of the limits of integration is needed. Consider first the lower limit, z_1. The photoelectrons originating at a level z below z' lose energy as they move upward along a field line according to Equation (5.3) with $\cos \alpha = 1$, and upon arrival at level z' have an energy W given by Equation (5.5); the lower limit, z_1, is such that $W = 0$ when $z = z_1$.

Two cases must be distinguished in a consideration of the upper limit, according as the level z' is above or below the escape level. If $z' \geq z_e$, then the upper limit is infinite. If $z' < z_e$, then z_2 is such that $W = 0$ when $z = z_2$ in the equation

$$W = E \left[1 + \frac{2H}{R(z)} \left\{ 1 - \exp \left[-(z' - z)/2H \right] \right\} \right]^{1/2} . \qquad (5.9)$$

This equation is identical to Equation (5.5) except for a change of sign that results because $\cos \alpha = -1$ for photoelectrons traveling in the downward direction.

The integration of Equation (5.8) will first be done for a level $z' < z_e$. After substitution for $\left| \frac{dW}{dz} \right|$ from Equation (5.3) and for W from Equations (5.5) and (5.9), Equation (5.8) takes the form

$$Q(z') = \frac{Eq(z_e)}{4R(z')}\left[\int_{z_1}^{z} \frac{\exp\left[-(z - z_e)/H\right] dz}{\left\{1 - \frac{2H}{R(z)}\left(1 - \exp\left[-(z' - z)/2H\right]\right)\right\}^{1/2}} + \right.$$

$$\left. \int_{z'}^{z_2} \frac{\exp\left[-(z - z_e)/H\right] dz}{\left\{1 + \frac{2H}{R(z)}\left(1 - \exp\left[-(z' - z)/2H\right]\right)\right\}^{1/2}}\right]. \qquad (5.10)$$

With the following substitutions:

$$\left.\begin{aligned} x^2 &= \exp(-z/2H) \\ A^2 &= \exp(-z_e/2H) \\ B^2 &= \exp(-z'/2H) \\ C^2 &= A^2 + B^2 \\ D^2 &= B^2 - A^2 \end{aligned}\right\}, \qquad (5.11)$$

Equation (5.10) can be put into the form

$$Q(z') = \frac{-HEq(z_e)}{A^3 R(z')}\left[\int_{C}^{B} \frac{x^3\, dx}{(C^2 - x^2)^{1/2}} + \int_{B}^{D} \frac{x^3\, dx}{(x^2 - D^2)^{1/2}}\right]. \qquad (5.12)$$

These integrals are now in a standard form found in the Dwight (1961) tables.

Upon evaluation of these integrals the expression reduces to

$$Q(z') = Eq(z'). \qquad (5.13)$$

This is the same heating rate as that which would apply if all the photoelectrons produced in unit volume at level z' were reduced to thermal energies within that same volume. Thus, the local heating rate prevails at all altitudes below the escape level.

This is not the case at altitudes above the escape level. From the definition of the variable x, it follows that the upper limit of the second integral in Equation (5.12) is zero. Upon evaluation of the integral for this upper limit there results for $Q(z')$ the expression

$$Q(z') = Eq(z') \left[1 + \frac{(1 - M^2)^{1/2}}{3} \left\{ \frac{1}{M^2} - 1 \right\} \right], \qquad (5.14)$$

where

$$M^2 = \exp \left[-(z' - z_e)/2H \right] .$$

The altitude dependence of this heat input function is represented by the fine broken line in Figure 5.2; the function has been normalized to unity at the escape level. Also shown here is the local heating rate, represented by the coarse broken line. At altitudes well above the escape level, the non-local heating rate has an exponential dependence upon altitude with scale height equal to that of the electron density distribution.

Because of the decrease of the rate of production of photoelectrons with altitude, there is a net upward flux of kinetic energy across any level z'; when z' becomes infinite, this is the escape flux, $F(\infty)$. At any level $z' \geq z_e$,

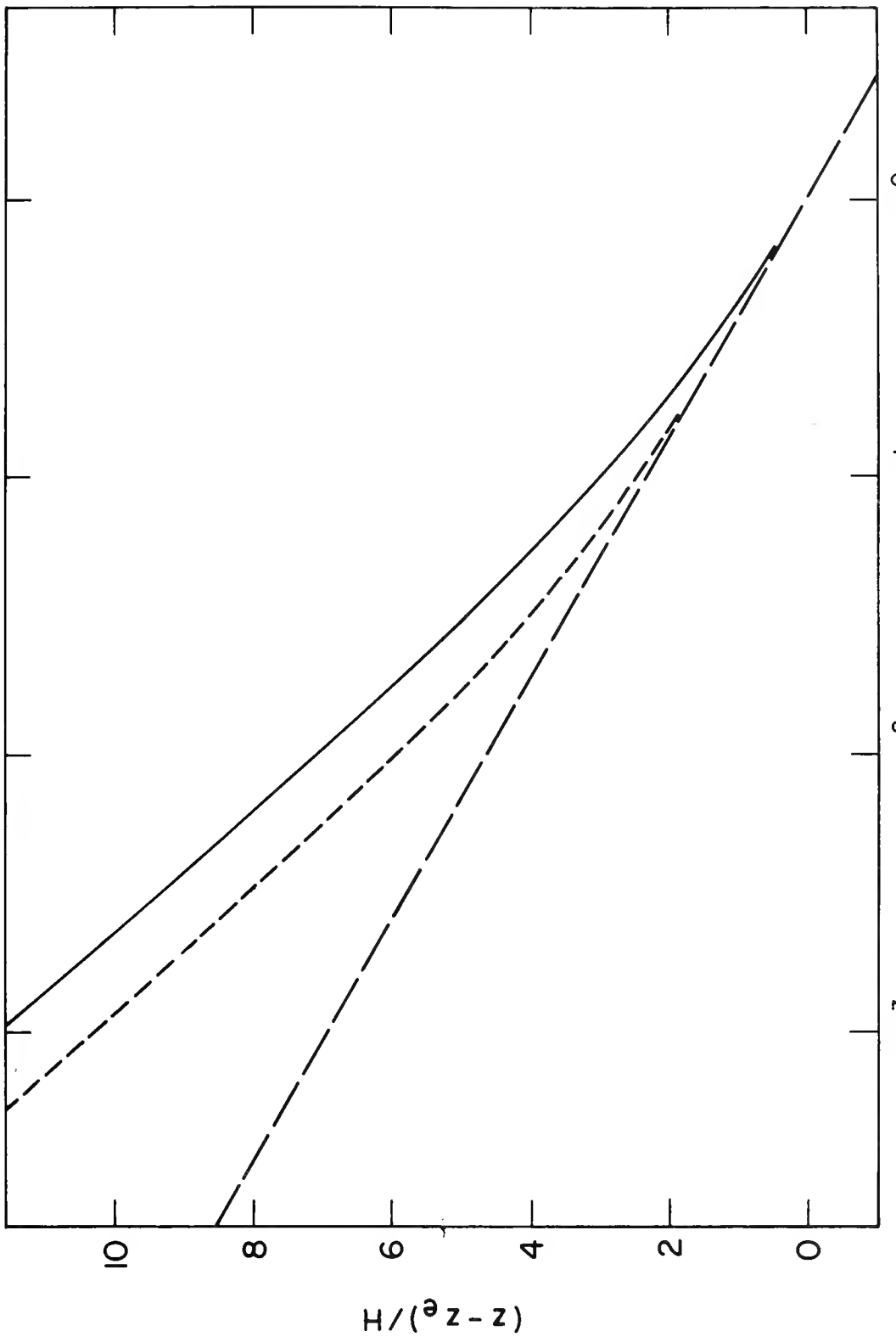

Figure 5.2 Altitude dependence of the local heating rate (coarse broken line) and the two non-local heating rates.

the net flux of kinetic energy is given by

$$F(z') = \int_{z_1}^{z'} W(z',z)\,\frac{q(z)}{2}\,dz - \int_{z'}^{\infty} W(z',z)\,\frac{q(z)}{2}\,dz. \qquad (5.15)$$

Upon substitution from Equations (5.5), (5.9) and 5.11), this relation takes the form

$$F(z') = \frac{-2HEq(z_e)}{A^5}\left[\int_C^B x^3(C^2 - x^2)^{1/2}\,dx - \int_B^0 x^3(x^2 - D^2)^{1/2}\,dx\right]. \qquad (5.16)$$

After evaluation of these integrals,

$$F(z') = \frac{8HEq(z_e)}{15}\left[1 - 1/2\ (1 - M^2)^{5/2}\right]. \qquad (5.17)$$

As $z' \longrightarrow z_e$, $M^2 \longrightarrow 1$, and the energy flux upward across the escape level is therefore a fraction 8/15 of the integrated rate of production of energy above the escape level. As $z' \longrightarrow \infty$, $M^2 \longrightarrow 0$, and the escape flux is therefore just half this amount. The remaining half is converted to heat and is responsible for the increase of the heating rate over that of local heating in the region above the escape level.

The escaping beam of photoelectrons is not monoenergetic, and it is of some interest to look at the energy spectrum. The integral relation for the escape flux follows from Equation (5.15):

$$F(\infty) = \lim_{z' \longrightarrow \infty} \int_{z_1}^{z'} W(z',z)\,\frac{q(z)}{2}\,dz. \qquad (5.18)$$

The energy at escape is given by

$$W = E\left[1 - \exp\left[-(z - z_e)/2H\right]\right]^{1/2}, \tag{5.19}$$

which follows from Equation (5.5) and the definition of R(z). To obtain the energy spectrum, it is necessary to eliminate the variable z in Equation (5.18) and express $F(\infty)$ in the form of an integral equation with the energy W as variable. Since $q(z) \propto \exp\left[-(z - z_e)/H\right]$, this change of variable can be accomplished by differentiating Equation (5.19) with respect to z and solving for $\exp\left[-(z - z_e)/2H\right]$. Substitution for q(z)dz in terms of this quantity may then be made in Equation (5.18), with the result that

$$F(\infty) = 2Hq(z_e)\int_0^E \frac{W^2}{E^2}\left\{1 - \frac{W^2}{E^2}\right\} dW \ . \tag{5.20}$$

By definition of the escape level, half of the photoelectrons produced above this level escape, and the number flux of escaping photoelectrons is therefore equal to $Hq(z_e)/2$. Division of both sides of Equation (5.20) by this quantity results in an expression for the mean energy of these photoelectrons:

$$\overline{W} = 4\int_0^E \frac{W^2}{E^2}\left(1 - \frac{W^2}{E^2}\right) dW. \tag{5.21}$$

It now follows that the quantity

$$\Psi(W) = \frac{4W}{E^2}\left(1 - \frac{W^2}{E^2}\right), \tag{5.22}$$

must be the energy distribution of these photoelectrons. Equation (5.22) is easily integrated over energy, and, as expected,

$$\int_0^E \Psi(W)\, dW = 1.$$

Integration of Equation (5.21) gives for the mean energy

$$\overline{W} = \frac{8}{15} E,$$

while integration of Equation (5.20) gives for the escape flux

$$F(\infty) = \frac{4}{15} HEq(z_e),$$

which is the same result obtained from Equation (5.17). The energy distribution is illustrated in Figure 5.3. It may be seen that it is somewhat skewed toward the lower energies, with a sharp cutoff at the maximum energy.

The energy carried by the photoelectrons that escape the F2-layer is the principal source of heat in the protonosphere. This is the subject of Chapter 6, but it can be pointed out at this time that the derivation of an expression for the heating rate in this region requires a knowledge of the energy distribution of the photoelectrons entering the protonosphere. This is given by Equation (5.22) for the electron beam approximation used in this section. The energy distribution remains unchanged except for a multiplicative factor that is a function of pitch angle when the pitch angle distribution of the photoelectrons is taken into account.

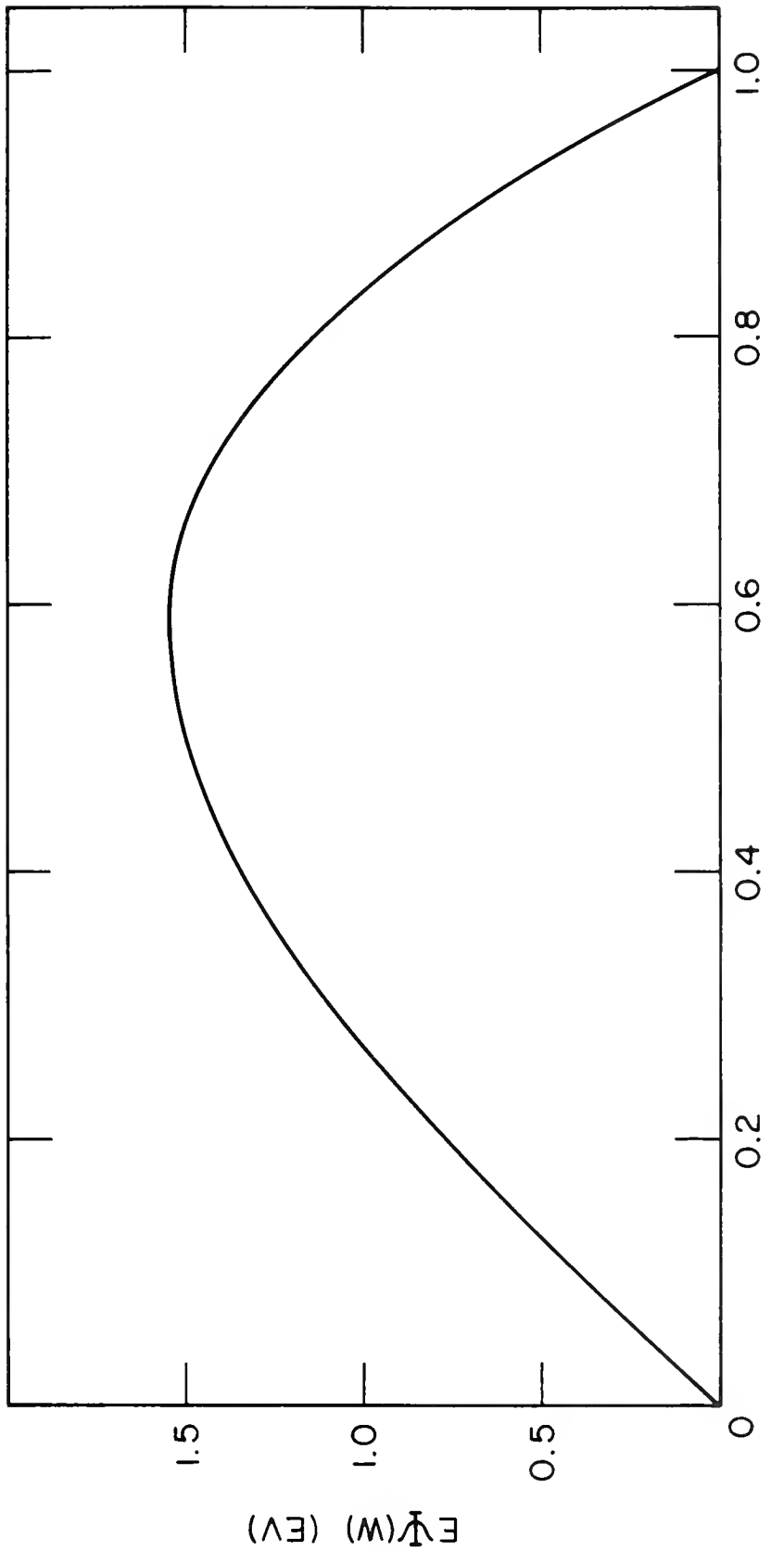

Figure 5.3 Energy distribution of escaping photoelectrons; values of $E\,\Psi(w)$ shown here are for E = 10 ev.

5.3.2 Non-Local Heating by Photoelectrons of Isotropic Pitch Angle Distribution

The method of treating the problem of the heating rate and the escape of photoelectrons that takes into account the pitch angle distribution of the photoelectrons is entirely analogous to the method of the preceding section. A minimum of detail will be given here, and frequent reference to the preceding section will be made.

Reference to Equations (5.3), (5.5), and (5.6) shows that both the range and the escape level are now functions of pitch angle. It is convenient however to retain the factor of cos α explicitly in the calculations, and in this section the terms range and escape level will, unless otherwise stated, refer to the range and escape level of a photoelectron with zero pitch angle.

The expression for the heating rate now contains an average over the pitch angle distribution. The analogue of Equation (5.8) is:

$$Q(z') = \frac{Kn(z')}{\sin I}\left[\left\{\int_0^{\pi/2}\frac{g(\alpha)d\alpha}{\cos\alpha}\int_{z_1}^{z'}\frac{q(z)dz}{W(z',z,\alpha)}\right\} - \left\{\int_{\pi/2}^{\alpha^*}\frac{g(\alpha)d\alpha}{\cos\alpha}\int_{z'}^{z_2}\frac{q(z)dz}{W(z',z,\alpha)}\right\} - \left\{\int_{\alpha^*}^{\pi}\frac{g(\alpha)d\alpha}{\cos\alpha}\int_{z'}^{\infty}\frac{q(z)dz}{W(z',z,\alpha)}\right\}\right], \quad (5.23)$$

where $g(\alpha)$ is the pitch angle distribution, and

$$W(z',z,\alpha) = E\left[1 - \frac{2H}{R(z)\cos\alpha}\left\{1 - \exp\left[-(z' - z)/2H\right]\right\}\right]^{1/2}. \quad (5.24)$$

The first term in curly brackets in Equation (5.23) represents the contribution from all photoelectrons originating below the level z'. The second term in curly brackets represents the contribution from all photoelectrons with pitch angles $\pi/2 < \alpha \leq \alpha^*$, where the photoelectron which originates at infinity with pitch angle α^* has zero energy upon reaching level z'. From Equation (5.24) it follows that

$$\cos \alpha^* = \lim_{z \to \infty} \frac{2H}{R(z)} \left\{ 1 - \exp \left[-(z' - z)/2H \right] \right\} .$$

This limit may be evaluated by substitution from the identity

$$R(z) = R(z_e) \exp \left[-(z_e - z)/2H \right]$$

to obtain

$$\cos \alpha^* = \lim_{z \to \infty} \left\{ \exp \left[-(z - z_e)/2H \right] - \exp \left[-(z' - z_e)/2H \right] \right\} ,$$

whence,

$$\cos \alpha^* = \exp \left[-(z' - z_e)/2H \right].$$

The contribution from the remaining photoelectrons which originate above the level z' and move downward is given by the third term in Equation (5.23). Note that as $z' \to z_e$, $\alpha^* \to \pi$, and this term therefore vanishes at the escape level.

Formally, the integrals over altitude in Equation (5.23) are identical to those of Equation (5.8), and when these are evaluated, Equation (5.23) becomes

$$Q(z') = Eq(z')\left[1 - \frac{2}{3}\int_{\alpha^*}^{\pi} (1 + M^2 \sec\alpha)^{3/2} \frac{\cos\alpha}{M^2}\, g(\alpha)\, d\alpha \right] . \qquad (5.25)$$

This shows the expected result that at and below the escape level, local heating prevails.

For the purpose of evaluating this integral, it is convenient to make the substitution

$$x^2 = - M^2/\cos\alpha \quad ,$$

and when the isotropic pitch angle distribution

$$g(\alpha) = \frac{\sin\alpha}{2}$$

is substituted, Equation (5.25) takes the form

$$Q(z') = Eq(z')\left[1 - \frac{2M^2}{3}\int_{1}^{M} \frac{(1 - x^2)^{3/2}}{x^5}\, dx \right] , \qquad (5.26)$$

which is again in the standard form found in the Dwight (1961) tables. Upon evaluation of this integral, the expression for the heating rate becomes

$$Q(z') = Eq(z')\left[1 + \frac{2}{3}(1 - M^2)^{1/2}\left\{ \frac{1}{4M^2} - \frac{5}{8} \right\} + \frac{M^2}{4} \log\left| 1 + \frac{(1 - M^2)^{1/2}}{M} \right| \right] . \qquad (5.27)$$

This heating function, normalized to unity at the escape level, is shown by the continuous line in Figure 5.2, where it may be compared with the heating function given by Equation (5.14). Again the heating rate has at altitudes well above the escape level an exponential dependence upon altitude with scale height equal to that of the electron density distribution.

The net upward flux of kinetic energy across any level $z' \geq z_e$ is given by

$$F(z') = \int_0^{\pi/2} g(\alpha)d\alpha \int_{z_1}^{z'} W(z',z,\alpha)q(z)dz - \int_{\pi/2}^{\alpha^*} g(\alpha)d\alpha \int_{z'}^{z_2} W(z',z,\alpha)q(z)dz$$

$$- \int_{\alpha^*}^{\pi/2} g(\alpha)d\alpha \int_{z'}^{\infty} W(z',z,\alpha)q(z)dz \ , \qquad (5.28)$$

which is the analogue of Equation (5.15). The integrals over altitude are formally the same as those of Equation (5.15), and upon evaluation of these integrals, Equation (5.28) becomes

$$F(z') = \frac{8}{45} HEq(z_e) \left[1 - 3 \int_{\alpha^*}^{\pi} \cos^2\alpha \ (1 + M^2 \sec\alpha)^{5/2} \ g(\alpha) \ d\alpha \right] .$$

As $z' \rightarrow z_e$, $\alpha^* \rightarrow \pi$ and the energy flux across the escape level is therefore a fraction 8/45 of the integrated rate of production of energy above the escape level. As $z' \rightarrow \infty$, $\alpha^* \rightarrow \pi$, and also $M^2 \rightarrow 0$, so that

$$3\int_{\alpha_*}^{\pi} \cos^2\alpha\,(1 + M^2 \sec\alpha)^{5/2}\,\frac{\sin\alpha}{2}\,d\alpha \rightarrow \frac{3}{2}\int_{\pi/2}^{\pi} \cos^2\alpha \sin\alpha\, d\alpha = \frac{1}{2}\,,$$

which shows that again the escape flux is exactly half the flux of energy through the escape level; the remaining half goes into heating the ionosphere above the escape level. Comparison with Equation (5.17) and the discussion following that equation shows that the escape flux is a factor of three lower in the present case.

In contrast to the results of Section (5.3.2), the results of the present treatment show that the topside ionosphere is less permeable to a flux of photoelectrons that have velocity components perpendicular to the magnetic field. The escape flux is cut down (by a factor of three for an isotropic pitch angle distribution) and a higher rate of heating is found.

There remains the question of the energy distribution of the escaping photoelectrons. This can be found from Equation (5.28) by the exact same method used earlier in the simplified electron beam treatment: elimination of the variable z' in favor of W. When this is done, Equation (5.28) in the limit as $z' \longrightarrow \infty$ becomes

$$F(\infty) = 4Hq(z_e)\int_0^{\pi/2} g(\alpha)\cos^2\alpha\, d\alpha \int_0^{E} \frac{W^2}{E^2}\left\{1 - \frac{W^2}{E^2}\right\} dW\,.$$

From comparison of this with Equations (5.20) and (5.22) it follows that the quantity

$$\Psi(W,\alpha) = 6g(\alpha)\cos^2\alpha\left\{\frac{4W}{E^2}\left(1 - \frac{W^2}{E^2}\right)\right\}, \qquad (5.29)$$

is the energy distribution of escaping photoelectrons in the sense that $\Psi(W,\alpha)dW\,d\alpha$ is the fraction escaping with pitch angles in the range $d\alpha$ about α and energies in the range dW about W.

5.4 Specification of the Heat Input Function

The physics of the heating of the electron gas has now been examined in sufficient detail that the practical side of the problem may be approached. It is the purpose of this section to formulate a heat input function that can be used in the calculation of the electron and ion temperature profiles.

The most comprehensive treatment of the local heating problem appears to be that of Dalgarno, et al. (1963). The large number of detailed calculations required in the treatment is apparent from the discussion of Section 5.2. This, coupled with the fact that no better information on either the distribution of the solar flux or the relevant cross sections is now available, has made a re-analysis seem undesirable. Accordingly the $Q(z')$ that is roughly the mean of the four heat input functions derived by Dalgarno, et al. has been adopted up to an altitude of 300 km. This $Q(z')$ is the heat input function at the time of the solar flux measurements; it must be scaled according to the procedure mentioned in Section 5.2 to obtain the $Q(z')$ at another epoch of the solar cycle.

The non-local heating rate is given by Equation (5.27). However, this expression was derived from consideration of photoelectrons all ejected with the same energy, E. Strictly speaking, the non-local heating rate is obtained only after the quantity defined by Equation (5.27) is averaged over the spectrum of initial energy. In view of the difficulty of such a calculation and the previously mentioned uncertainty in the distribution of the solar flux and in

the knowledge of the ionization cross sections, another procedure has been adopted: the escape level z_e has been chosen to be the 300 km level--in accordance with the original suggestion of Hanson (1963)--and $Q(z_e)$ has been set equal to the mean of the four values of Dalgarno, et al. at this level.

The heat input function $Q(z')$ has now been specified. Figure 5.4 shows this function scaled to represent the situation at the minimum of the solar cycle. The local heating rate has been extrapolated above the 300 km level as an exponential function with scale height equal to that of the ionizable constituent in this region (atomic oxygen) and is shown here for comparison.

It was pointed out in Section 5.3 that a substantial number of the photoelectrons produced above the escape level escape the ionosphere entirely and spiral up into the protonosphere. The ultimate fate of these photoelectrons will be considered in detail in Chapter 6. Suffice it to say that the photoelectrons of small pitch angle have sufficient range to penetrate the conjugate ionosphere. This is an additional source of heating that is not included in the heating function that has just been adopted. However, for reasons that will be stated later, the electron and ion temperatures are sufficiently insensitive to the heating rate above the escape level that this omission is of little consequence.

5.5 Cooling of the Electron Gas

The electron gas may cool through excitation of vibrational or rotational levels of molecular oxygen and molecular nitrogen. A somewhat less efficient but non-negligible method is energy transfer through elastic collisions with the neutral constituents. There is still another method of cooling: energy transfer via coulomb interactions with the ion gas, and subsequent cooling of

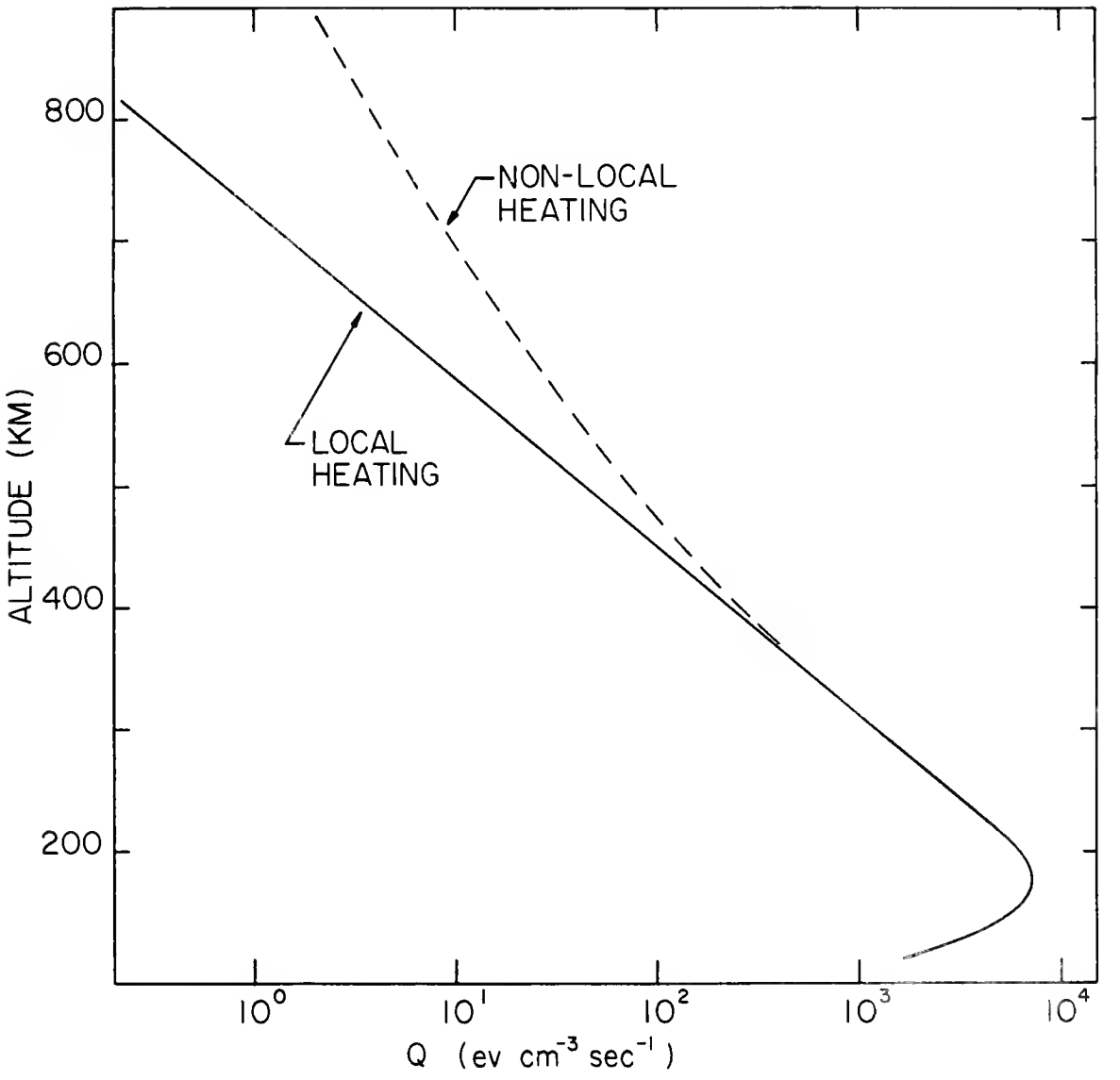

Figure 5.4 The rate of heat input to the ambient electron gas. Values at and below 300 km are roughly the mean of the four heat input functions derived by Dalgarno et al. (1963). The local heating rate above 300 km is obtained by assuming an exponential decrease with scale height of atomic oxygen at a temperature of 1000°K. The non-local heating rate takes into account an additional contribution from photoelectrons originating below 300 km.

the ion gas through ion-neutral elastic collisions. The ion gas here serves in the capacity of a coolant. However, as a consequence of the decrease of the cross section for energy transfer via coulomb collisions with increase of electron temperature, there is a limit to the rate at which the electron gas may be cooled by this process.

5.5.1 Cooling by the Neutral Atmosphere

The three primary constituents of the upper atmosphere (O, O_2, N_2) are all effective in cooling the electron gas. The rate L_n at which unit volume of the electron gas transfers thermal energy directly to these neutral constituents can be written as the sum of three terms:

$$L_n = L_O + L_{O_2} + L_{N_2}. \tag{5.30}$$

Each of these terms is in general a function of the electron temperature T_e, the neutral gas temperature T_n, and the product of the number density of electrons and the number density n(X) of the constituent X.

The rate of heat loss through elastic collisions with atomic oxygen is given (in c.g.s. units) by

$$L_O = \frac{3}{2} k n_e \left[1.0 \times 10^{-14} T_e^{1/2} \right] n(O) (T_e - T_n), \tag{5.31}$$

where k is Boltzmann's constant. The coefficient in the square brackets has been taken from Dalgarno, et al. (1963). The rate of loss to molecular oxygen is given by

$$L_{O_2} = \frac{3}{2} k n_e \left[4.0 \times 10^{-14} T_e - 8.0 \times 10^{-12} \right] n(O_2) (T_e - T_n), \tag{5.32}$$

which represents loss of energy due primarily to collisional excitation of the rotational levels of O_2, with a small contribution from excitation of vibrational states. The coefficient in the square brackets has been taken from Hanson (1963). For loss to molecular nitrogen the expression

$$L_{N_2} = \frac{3}{2} k n_e \left[7.6 \times 10^{-16} T_e + \left\{ 1.2 \times 10^{-11} - 5.6 \times 10^{-15} T_n \right\} \right] n(N_2)(T_e - T_n), \quad (5.33)$$

has been adopted. The first term in the square brackets represents loss due to elastic collisions; the coefficient has been taken from Dalgarno, et al., (1963). The second term, grouped within the curly brackets, has been obtained by fitting the theoretical result of Dalgarno and Moffett (1962) with a linear expression. The process represented by this term is primarily one of excitation of the rotational levels of N_2.

A more extensive discussion of these collision processes, and complete references to the measurements from which the cross sectional data were taken is contained in the references cited.

5.5.2 Cooling by the Ion Gas

The electron gas may also lose energy to the ion gas through coulomb interactions at a rate

$$L_{ei} = 4.82 \times 10^{-7} n_e^2 (1 + 3b) (T_e - T_i) T_e^{-3/2}, \quad (5.34)$$

where T_i is the ion temperature and b is the ratio of number density of helium ions to number density of electrons. This expression was first derived by Hanson and Johnson (1961) for the case of no helium ion content. The ion gas

then transfers heat to the neutral gas at a rate

$$L_{in} = \left[\left\{ 3.3 \times 10^{-14} \; b + 1.8 \times 10^{-13} \; (1 - b) \right\} n \; (O) \right.$$

$$\left. + \left\{ 1.2 \times 10^{-13} \; b + 3.6 \times 10^{-14} \; (1 - b) \right\} n \; (He) \right] n_e (T_i - T_n), \qquad (5.35)$$

where n(He) is the number density of helium--an expression due to Willmore (1964). The numerical coefficient for transfer from oxygen ions to atomic oxygen has been taken from Hanson (1963).

Heat transfer directly to the neutral atmosphere is in general the dominant mechanism for cooling the electron gas below about 200 km, while heat transfer by way of the ion gas becomes increasingly important at higher altitudes. Consider a region of the ionosphere well above this transition level and suppose that the rate of heat input is locally balanced by the rate of heat loss to the ion gas. Then Q may be substituted on the left hand side of Equation (5.34), and the behavior of the difference between ion and electron temperature is as shown in Figure 5.5.

Now the ion temperature is never less than the neutral gas temperature, which above 300 km is never less than about $1000^{o}K$ during the day. The curve with $T_i = 1000^{o}K$ is therefore an upper limit in the following sense: for low values of electron density and/or high values of Q, the value of Q/n_e^2 may be sufficiently high that it falls above this curve; this means that cooling to the ions cannot locally balance the heat input Q no matter how large the difference between electron and ion temperature. This has been termed the "runaway electron temperature condition" (Hanson and Johnson, 1961), and when

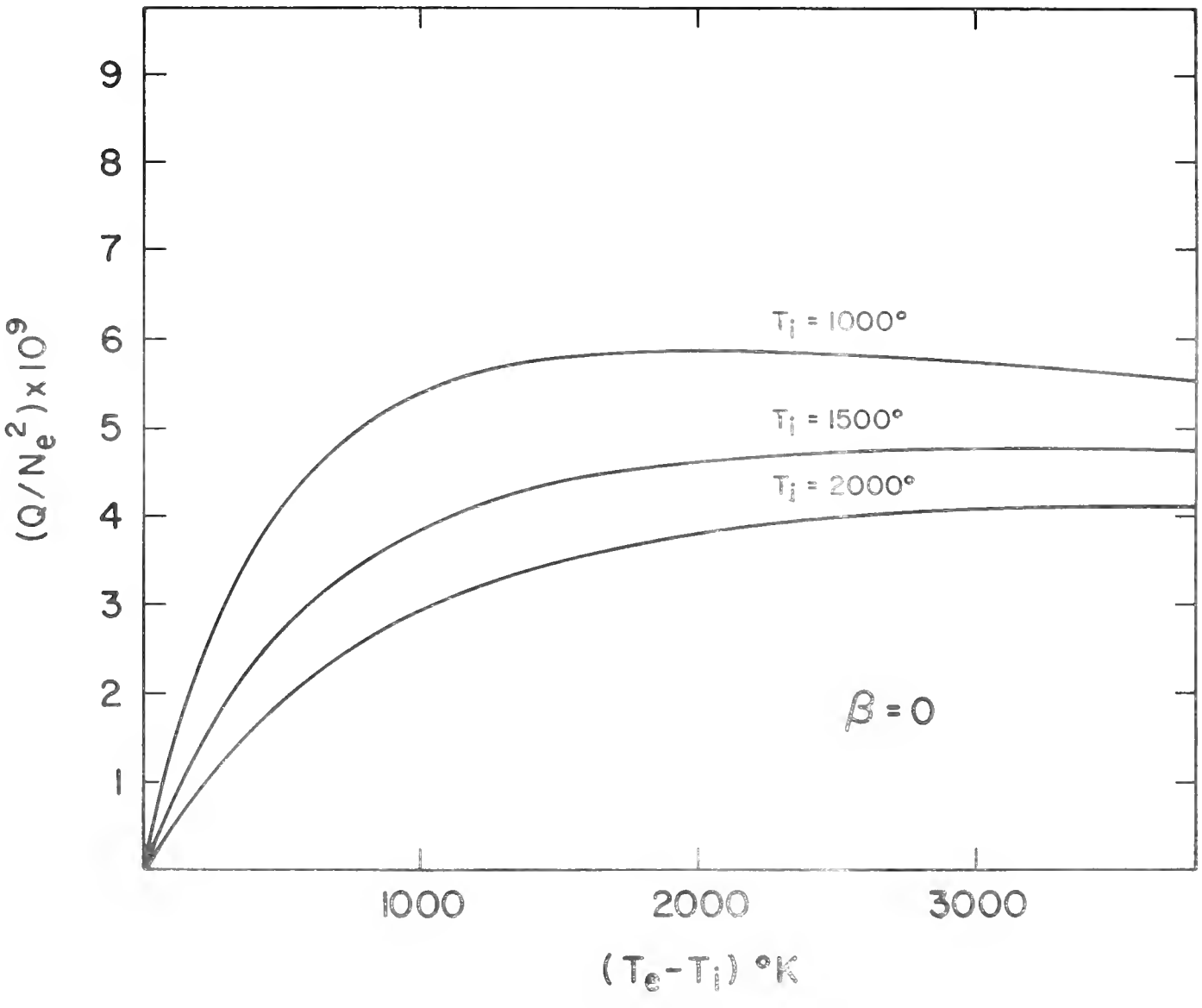

Figure 5.5 Difference between electron and ion temperatures required for rate of heat input to be locally balanced by rate of heat loss to the ion gas. The rate of heat loss depends on n_e^2 (Equation 5.34), and if $Q/n_e^2 > 6\times10^9$ ev cm^3/sec the balance cannot be achieved and electron temperature "runaway" occurs.

this condition sets in, the electron temperature will rise until the combined effect of thermal conduction and heat loss direct to the neutral atmosphere balance the excess Q.

5.6 Thermal Conduction

The thermal conductivity of a gas of charged particles is dependent upon the inverse square root of the mass of the particle (Chapman and Cowling, 1958; Spitzer, 1962). The thermal conductivity of the ion gas is therefore quite small compared to that of the electron gas, and thermal conduction by the ion gas may be neglected as a mechanism of heat transport. The rate of heat input to the ion gas is therefore locally equal to the rate of heat loss, and the equation which results from the equality of the right-hand sides of Equations (5.34) and (5.35) may be solved explicitly for the ion temperature in terms of the electron and neutral gas temperatures.

The equation governing the altitude dependence of electron temperature is the steady-state heat conduction equation:

$$\lambda \sin^2 I \frac{dT_e}{dz} = \int_z^\infty (Q - L)\, dz, \qquad (5.36)$$

where λ is the thermal conductivity of the electron gas along a magnetic field line and I is the magnetic dip angle; the rate of heat loss L is given by the sum of Equations (5.30) and (5.34). According to the expressions given by Spitzer (1962):

$$\lambda = 7.7 \times 10^5 T_e^{5/2} \text{ ev/cm sec deg.}$$

In order to illustrate some of the effects of thermal conduction, we consider three situations that are probably found within the ionosphere:

(i) Overheated electron gas

Suppose that the heat input and electron density are such that the runaway condition is present at all altitudes. There is then an excess of heat input over heat loss (termed heat excess below) above about 200 km, and very high electron temperatures will occur, due to the inefficacy of the ion gas coolant. The electron gas attains what we will call an overheated condition in this region. Because of the strong temperature dependence of the thermal conductivity, it is characteristic of an overheated electron gas to be very nearly isothermal. On the other hand, the increase of neutral concentration with decrease in altitude below the 200 km level implies the existence of a heat sink in the lower atmosphere. The result is that a strong electron temperature gradient is set up between the overheated region and the region of the heat sink. In the steady state, the temperature of the overheated electron gas is set by the requirement that this gradient be sufficient to conduct downward the integrated heat excess above 200 km.

(ii) Topside heat sink

Now, because ion and electron temperatures exceed that of the neutral atmosphere, the quantity Q/n_e^2 may decrease with altitude through the topside ionosphere. This raises the possibility that the runaway condition may be absent at very high altitudes, thus

providing what we may describe as a topside heat sink. However, the thermal conductivity of the overheated electron gas is so high that even a small decrease of electron temperature with altitude is attended by a heat flux that is sufficient to exceed the capacity of this sink. Thus, the overheated electron gas cannot be cooled by a topside heat sink.

(iii) The parasitic condition in an overheated electron gas

Consider the case in which the runaway condition exists only below 250 km. There are now two heat sinks: one below 200 km that is provided by the high concentration of neutrals and one above 250 km provided by the ion gas coolant. This situation leads to an electron temperature profile that exhibits a maximum at about 200 km. The temperature at the maximum is controlled by the concentration of neutrals at that altitude rather than by thermal conduction. However, this temperature is high enough to overheat the electron gas above the temperature maximum, so that strong temperature gradients are impossible at higher altitudes. This means that the high altitude electron gas is parasitic upon the runaway region (below 250 km), in the sense that heat is readily conducted from below to an extent which by far exceeds the capacity of the topside heat sink.

It will subsequently be demonstrated that the sunspot minimum and the sunspot maximum topside electron gases are both overheated; the sunspot minimum electron gas because the runaway condition persists at all altitudes, and the sunspot maximum electron gas because it is parasitic upon the runaway region below 250 km.

5.7 Calculation of Ionospheric Temperature Profiles

Sufficient material has now been assembled that the calculation of electron and ion temperature profiles may proceed. The electron temperature is obtained by numerical integration of Equation (5.36). At each step of the integration the ion temperature is obtained from the equation that results from equality of the right-hand sides of Equations (5.34) and (5.35).

As pointed out in the introduction to this chapter, the techniques for measurement of electron and ion temperatures have been developed quite recently and most of the available temperature data are for a period of low solar activity. For this reason the solution has been carried out with a model atmosphere and model ionosphere representative of conditions near sunspot minimum. However, it was found in the course of the analysis that a clearer understanding of the reasons for the derived behavior of the temperature profiles could be attained by comparison with theoretical profiles calculated for the conditions which exist at sunspot maximum.

5.7.1 Model Atmosphere and Ionosphere

The temperature and number density of the neutral constituents are taken from the model atmospheres of Harris and Priester (1962) at time of local noon. Following the suggestion of Harris and Priester (1963), we have chosen the model S = 100 as most representative of conditions near the minimum of solar activity.

The sub-peak model ionosphere has been taken from an average of quiet day measurements of electron density recorded at Slough during September 1953 (Thomas, Haselgrove, and Robbins, 1957). Above the peak the electron density profile is smoothly joined to a model ionosphere consisting of oxygen ions,

helium ions, and electrons in diffusive equilibrium. The ratio of helium ion concentration to that of oxygen ions has been taken to be 5×10^{-2} at an altitude of 500 km. The electron temperature shows little altitude variation over the topside, but this is not the case with the ion temperature. In the region of the peak the ion and neutral gas temperatures are equal; at very great altitudes the ion temperature is equal to that of the electrons. More elaborate procedures are possible, but in view of the fact that the integration is started at a high altitude and carried downward, it is expedient to allow for this temperature variation by setting $T_i = (T_e + T_n)/2$ in the computation of the scale height of the ionization.

5.7.2 Method of Solution

Because of the heat sink provided by the high concentration of neutral constituents at low altitudes, the difference between electron and neutral gas temperatures must ultimately decrease with decreasing altitude. If the level where this difference can be considered negligibly small lies below the level where Q reaches its maximum value (about 180 km), then this difference will remain negligibly small at all lower altitudes. This provides boundary conditions for the solution of Equation (5.36): the electron temperature and its first derivative must have the same value as that for the neutral gas at some altitude below 180 km. What altitude this may be is not known *a priori*, however, and this makes it necessary to search for the solution by assuming initial values of electron temperature at some high altitude and integrating downward.

In practice, this searching process required a large number of trials, and for this reason, the integrations were performed by the IBM 7094 computer at the University of Illinois. Two values of electron temperature that might

reasonably be expected to bracket the correct initial value were chosen, and the integration was started with the mean of these two values. If this initial value was too low, the electron temperature profile crossed that of the neutral atmosphere at a steep angle. This type of behavior is illustrated by curve (A) of Figure 5.6. In such a case, the integration was started again with a value of electron temperature that was the mean of the previous initial value and the upper bracketing temperature. In general, this initial value proved to be too high, and the electron temperature diverged from that of the neutral atmosphere at lower altitudes, resulting in a profile like curve (C) in Figure 5.6. Then, this initial value and the preceeding one were taken as upper and lower bracketing temperatures, respectively, and the succeeding initial value was taken as the mean of these two. This process was continued until two successive initial values differed only in the ninth significant figure. The solution satisfying the boundary conditions exactly--curve (B) in Figure 5.6--was never obtained by this process, but in most cases the last two integrations provided electron temperature profiles that were identical except for small differences in the last few kilometers of the integration range.

The solutions to be described were obtained for a geomagnetic latitude of 40° ($I \approx 70^{\circ}$). The integrations were started at the 1000 km level, and it was assumed that the heat flux through this surface was zero. This is not strictly true, since the protonosphere is continually heated during the daylight hours by the photoelectrons that escape the F2-layer and, in the steady-state situation, this heat must be removed by thermal conduction downward. An estimate of the flux of energy due to the escaping photoelectrons shows that

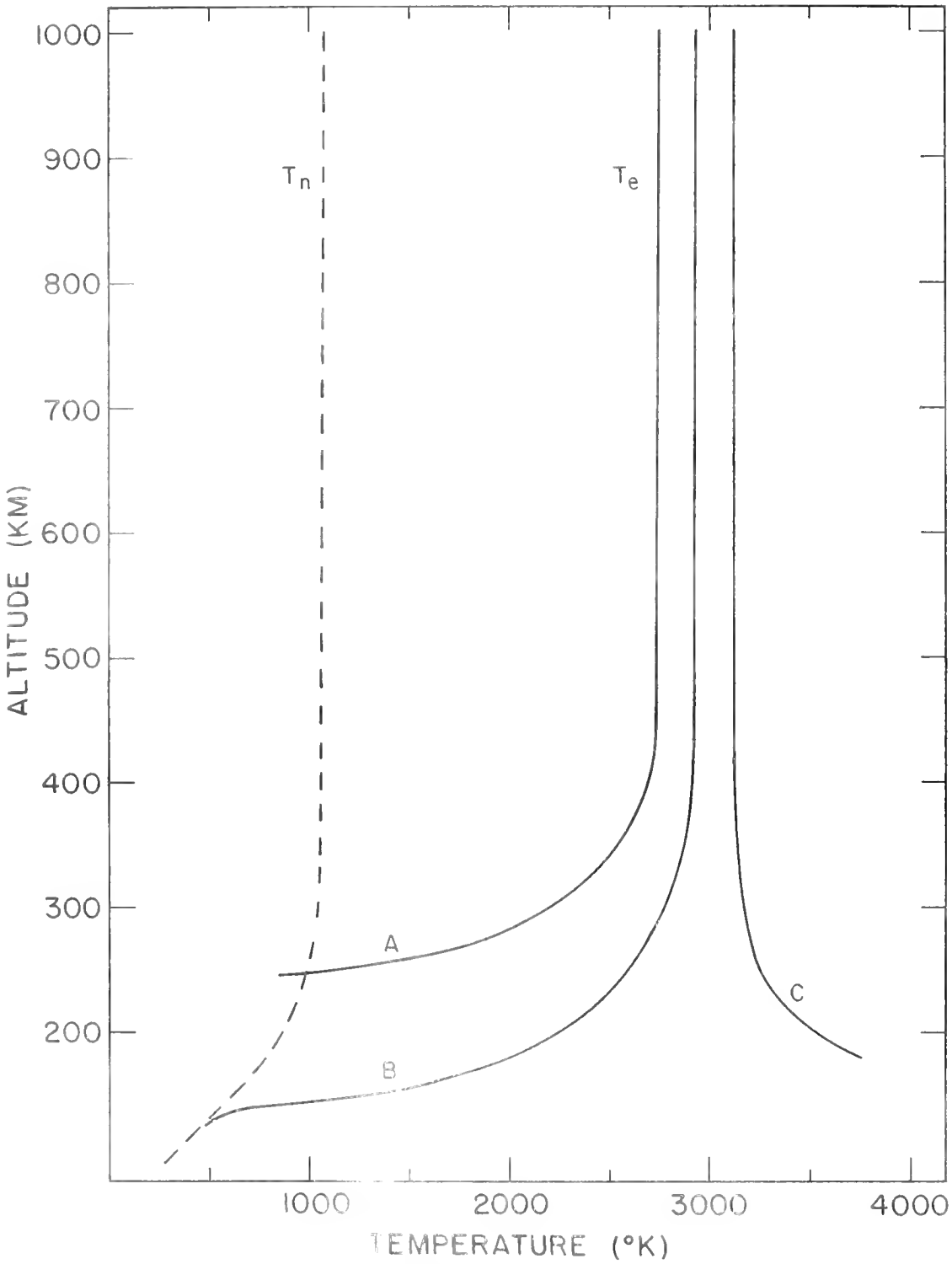

Figure 5.6 Behavior of the trial integrations of Equation (5.36). Profile of the type (A) results when the assumed initial value of T_e at 1000 km is too low, while profile of the type (C) results when this value is too high. Profile (B) is the solution satisfying the lower boundary conditions.

the flow of heat downward through the 1000 km surface is probably not in excess of 3×10^{8} ev cm^{-2} sec^{-1}. Some integrations were performed with this heat flux incorporated, but the change in the temperature profiles was not significant enough to warrant further discussion.

5.8 Results and Discussion

5.8.1 The Sunspot Minimum Model

The calculated temperature profiles for sunspot minimum are shown in Figure 5.7. Evaluation of the ratio Q/n_e^2 shows that the runaway condition persists at all altitudes in this model. Thus, there is an excess of heat input over heat loss in the altitude range over 200 km, and the electron temperature profile is set by the requirement that this heat be conducted downward to lower altitudes, where it can be transferred directly to the neutral gas.

The ion temperature is close to that of the neutral atmosphere below about 500 km, but with further increase of altitude it converges to that of the electron temperature. This is a direct consequence of the steady decrease of the ratio of neutral density to ion density with altitude; at the higher altitudes, the ions must lose thermal contact with the neutral atmosphere.

The large difference between the local and non-local heating rates can be seen in Figure 5.4. However, this difference does not affect the electron temperature profile. This is seen in Figure 5.8. The local heating rate is sufficiently low that the runaway condition is absent above about 500 km, and a topside heat sink is therefore present. However, an upward heat flux that is less than one per cent of the rate of accumulation of excess heat in the region between 200 and 500 km is sufficient to exceed the capacity of this

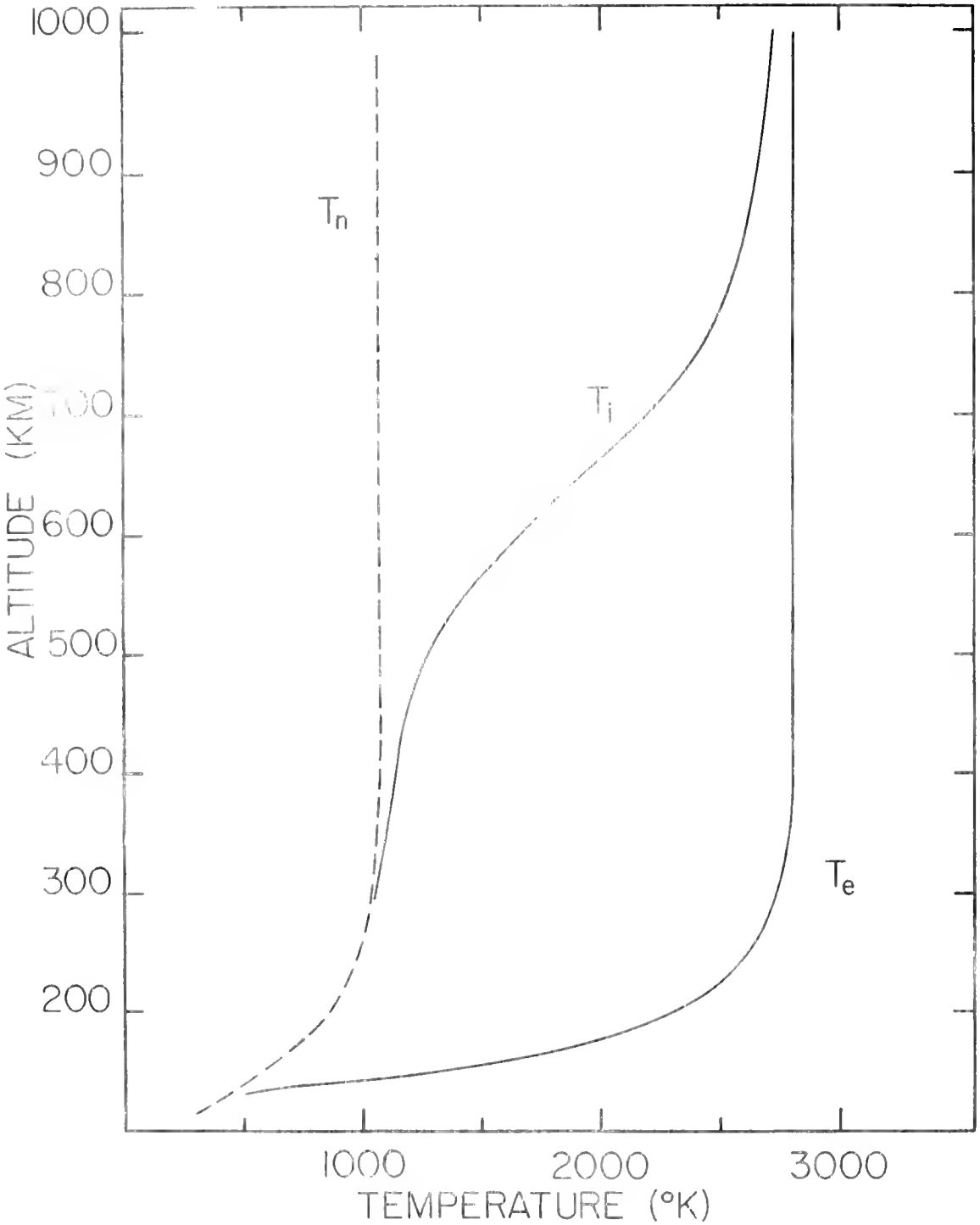

Figure 5.7 Ion and electron temperature profiles at sunspot minimum. Temperature of the neutral atmosphere (broken line) is taken from Harris-Priester model S=100 at 1200 local time.

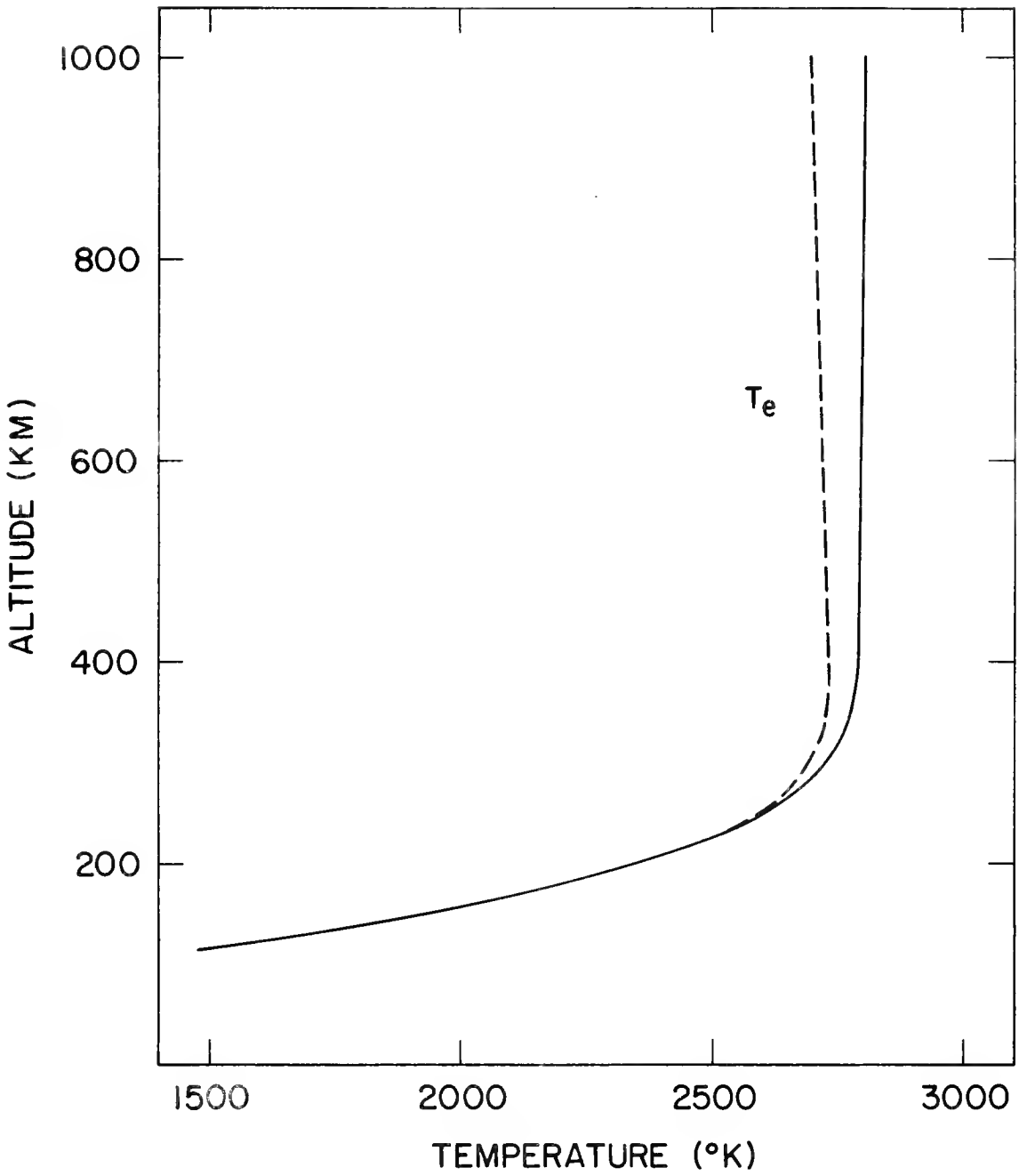

Figure 5.8 The effect on the electron temperature profile of a change in Q at hight altitudes. Continuous line represents same profile as that in Figure 5.7, while broken line represents profile that results when local heating is assumed above 300 km.

heat sink. Since the balance must be conducted downward, the electron gas in this region remains in an overheated condition. A very slight negative temperature gradient is then required to maintain the heat flow into the topside heat sink, and a nearly isothermal condition prevails.

The only way in which the heating rate above 500 km could appreciably affect the electron temperature at that altitude would be to lower Q or increase the electron density sufficiently between 200 and 500 km to prevent runaway in that region. It therefore seems unlikely that the inclusion of heating due to photoelectrons originating in the conjugate hemisphere would have much effect, since this would serve only to increase Q.

The reduction of incoherent backscatter spectra to obtain ion and electron temperatures requires some assumption about the relative abundance of ionic species at the level from which the signal is returned. For example, the electron temperature at 720 km deduced, under the assumption that the ion composition there is 20 per cent helium and 80 per cent oxygen, is some 15 per cent lower than that deduced under the assumption of oxygen ions only (Evans, 1964). However, it was shown above that the electron temperature is insensitive to the rate of heat input above the 500 km level, and also to the rate of heat loss above this level. Thus, even though the loss rate depends strongly on the ion composition (see Equation 5.34), one might surmise that the relative abundance of helium ions would not greatly affect the electron temperature deduced. This is confirmed by Figure 5.9, which shows that the response of both electron and ion temperature to change in helium ion content is slight.

5.8.2 Scaling of Parameters

Certain consequences of the theory are best illustrated by scaling the various physical quantities and solving the conduction equation numerically

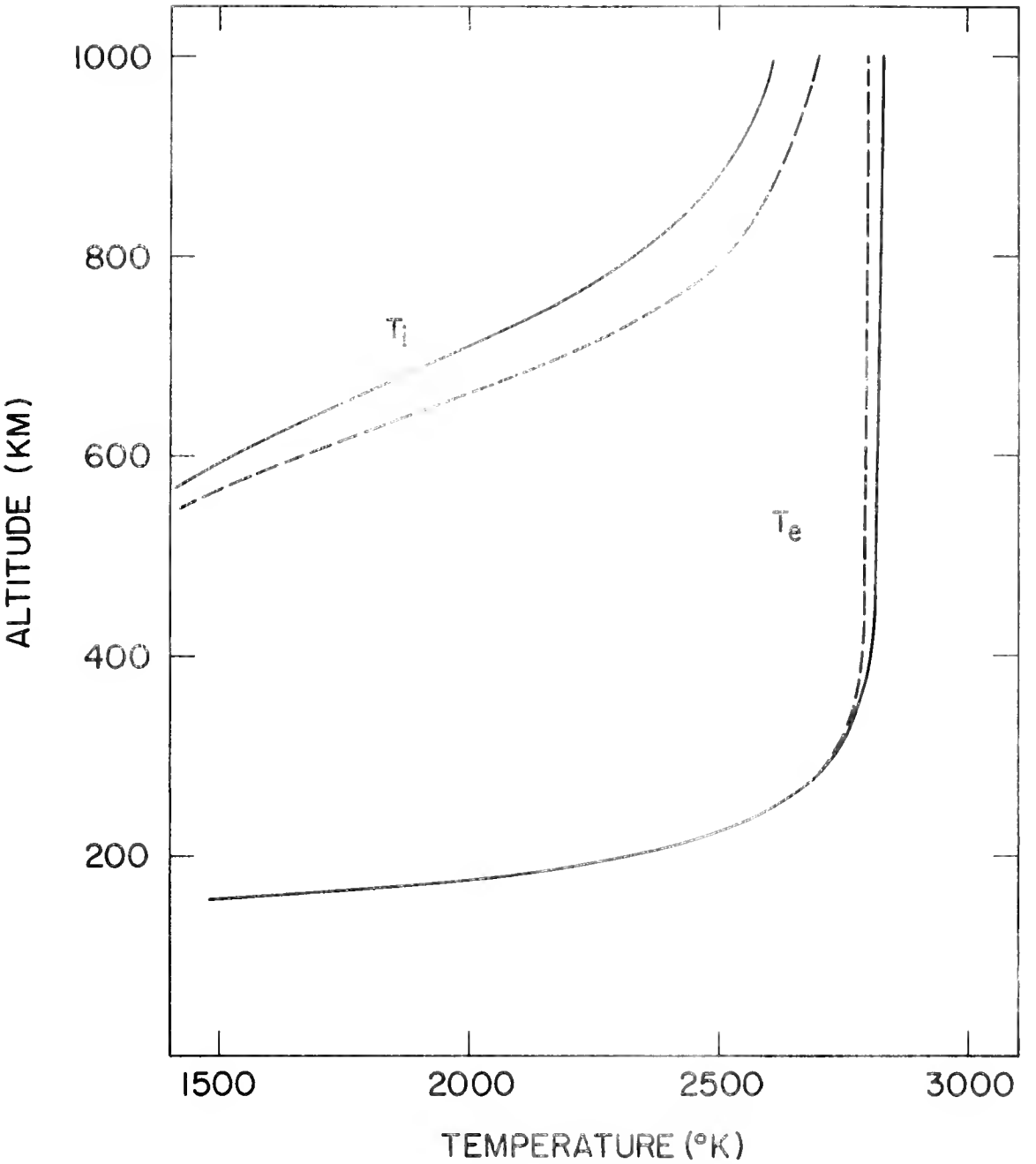

Figure 5.9 The effect on electron and ion temperature profiles of a change in relative abundance of helium ions. Profiles represented by the continuous line are obtained for 20% He^+ at 700 km; those represented by the broken line are obtained for 0.1% He^+ at 700 km.

for different combinations of the scale factors. In particular, the transition from sunspot minimum to sunspot maximum can be approximated in this way. Height independent scale factors have therefore been inserted into the electron density, the rate of heat input, and the combined neutral density; these are denoted by S_e, S_Q, and S_n, respectively.

Figure 5.10 shows electron temperature profiles obtained by scaling the rate of heat input, Q. This illustrates the sensitivity of the sunspot minimum electron temperature to a change in the incoming solar flux: there is an almost linear relation between T_e and S_Q above 300 km. Considering the uncertainty in the measurements of the solar flux and the nature of the assumptions that must be made in the calculation of Q, it is possible that this quantity may err by a factor of two. Figure 5.10 therefore indicates the possible uncertainty in the theoretical electron temperature at sunspot minimum.

In order to illustrate the transition to sunspot maximum conditions, the electron density has been scaled up by a factor of 5; the behavior of the temperature profiles is illustrated in Figure 5.11. Consider the cases with Q scale factors of 0.5 and unity. Calculation of Q/n_e^2 has shown that the runaway condition is not present at any altitude in these two cases, and inspection of Figure 5.11 shows that the depature from thermal equilibrium is small. But with further increase of Q, the runaway condition sets in below 250 km, and this is responsible for the increase of electron temperature between 250 and 200 km. As discussed in Section 5.6, the electron temperature rises until the excess heat is removed by heat loss direct to the neutrals. This is further illustrated in Figure 5.12: the neutral atmosphere serves as a moderator of electron temperature runaway in this altitude range. As a result, thermal conduction is a

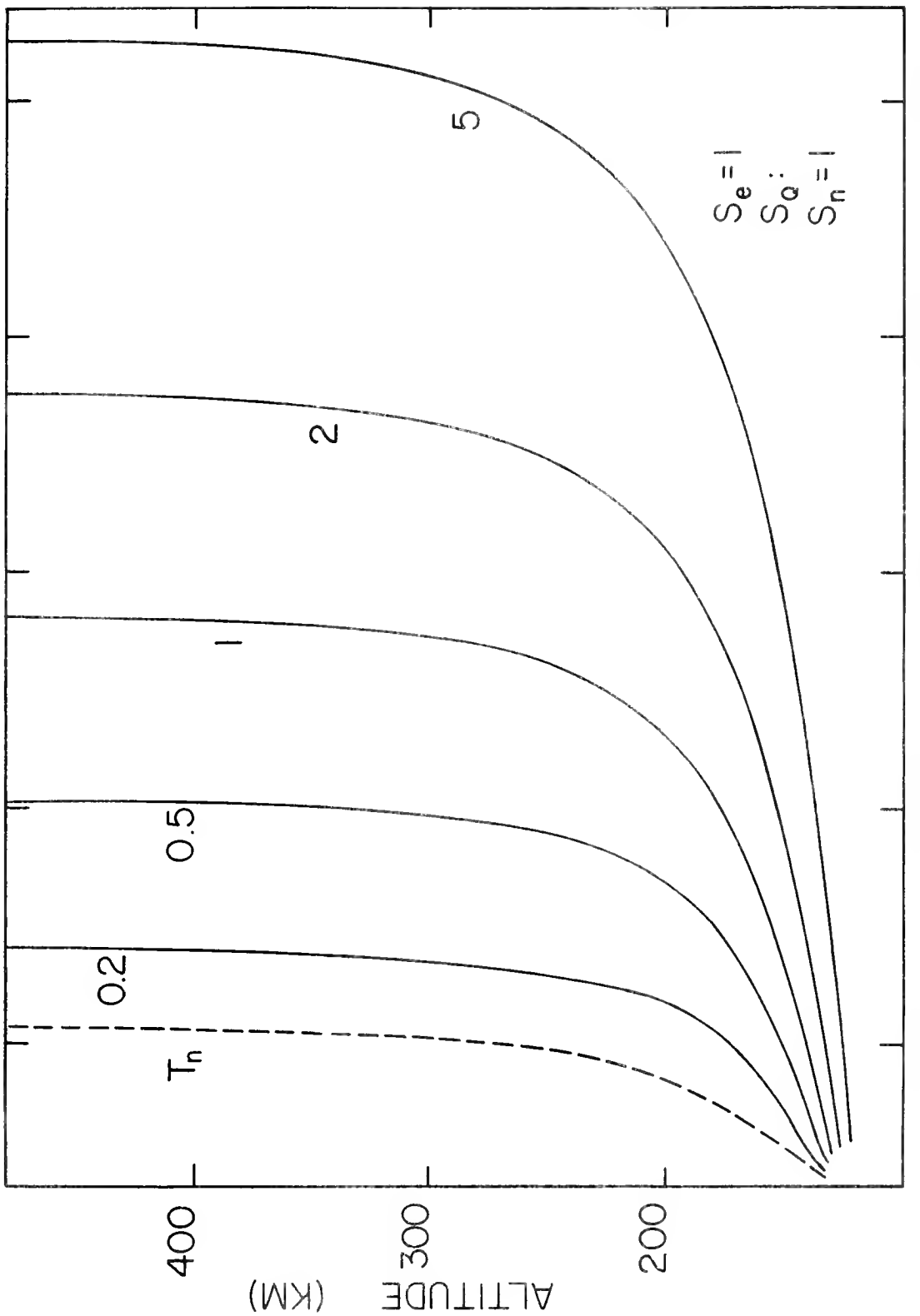

Figure 5.10 The effect on the electron temperature profile of scaling Q by the factor shown beside each curve.

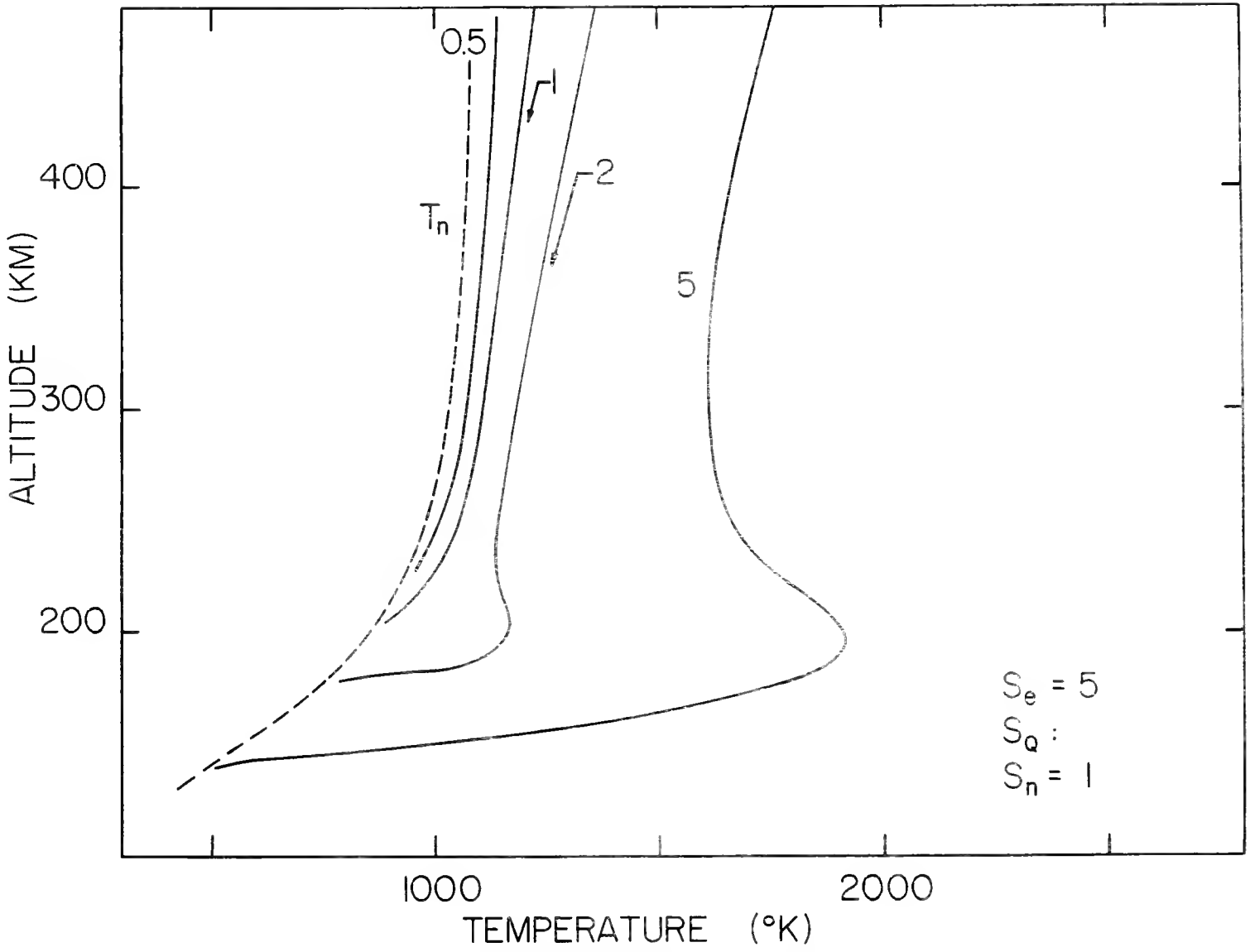

Figure 5.11 The electron density of the sunspot minimum model has been scaled up by a factor of 5 to simulate electron density at sunspot maximum, and the effect on the electron temperature profiles of scaling Q by the factor shown beside each curve is as seen here.

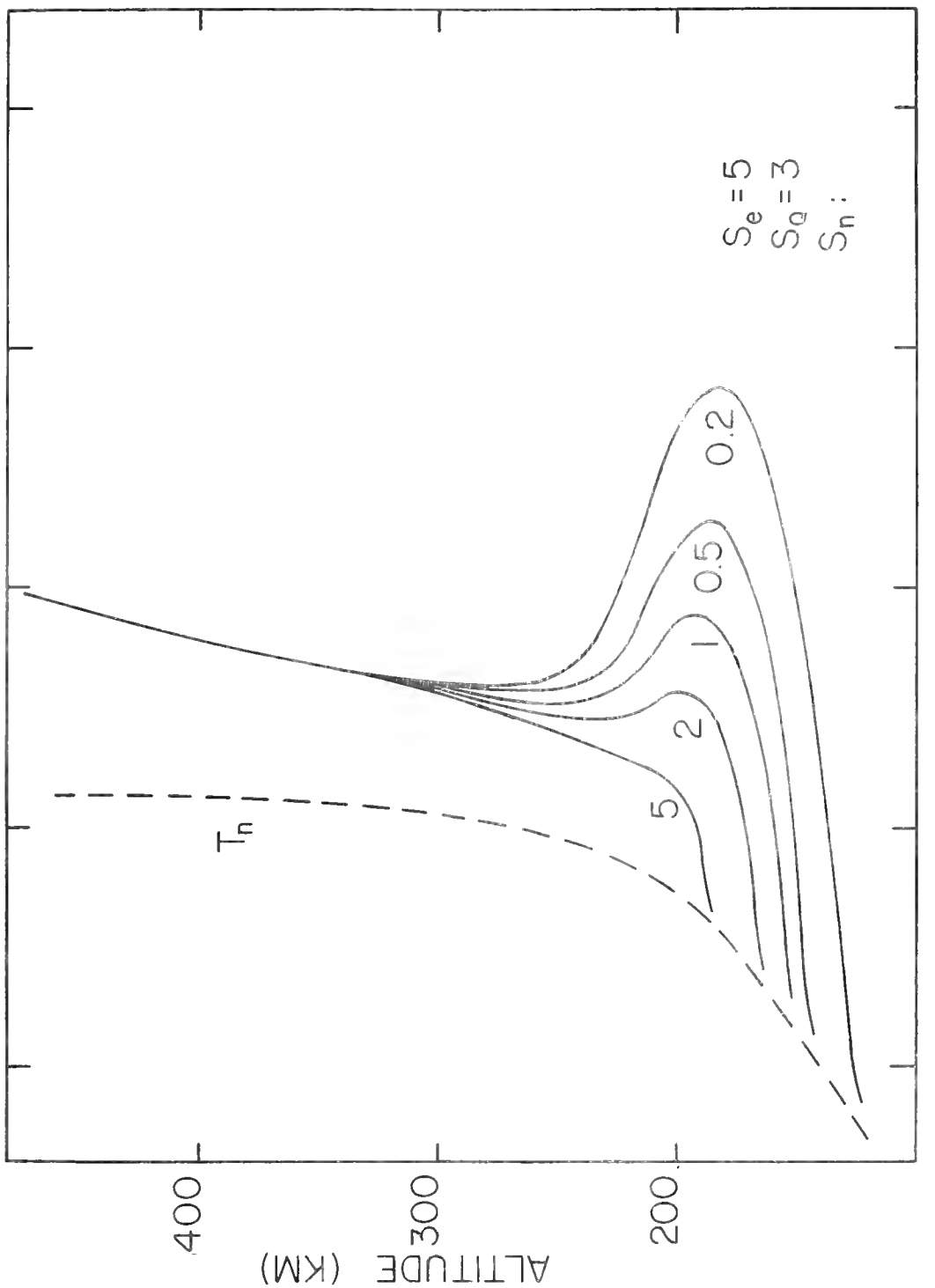

Figure 5.12 Illustrates the control of electron temperature by the neutral constituents when the runaway condition is present only below 250 km. The electron density and the Q of the sunspot minimum model have been scaled by factors of 5 and 3, respectively.

second order influence and these profiles have the shape characteristic of those predicted by the theory which ignores the thermal conductivity of the ionosphere.

A model that closely approximates sunspot maximum conditions is obtained by scaling the electron density and the Q of the sunspot minimum model by factors of 5 and 3, respectively. As shown in Figure 5.12, the resultant electron temperature profile has a pronounced maximum near 200 km. This leads to the suggestion that profiles derived by Hanson (1963) and by Dalgarno, et al. (1963) might be representative of sunspot maximum conditions. In order to strengthen this conclusion, an integration was carried out using a model ionosphere and a Harris-Priester model atmosphere representative of conditions at sunspot maximum. Electron and ion temperature profiles so obtained appear in Figure 5.13. The runaway condition is again present only in the region below 250 km. The pronounced maximum is present near 225 km, and the electron temperature here exceeds that of the neutral atmosphere by almost a factor of two. A comparison with the profiles of Dalgarno, et al. (1963) is shown in Figure 5.14. The hatched area covers the range of temperature profiles obtained by these investigators for the four assumed heat input functions. The heat input used in the sunspot maximum model was close to the largest of these. The principal differences between the profiles can be attributed to the effect of thermal conduction in destroying steep temperature gradients.

It is interesting to note that for the artificial models shown in Figure 5.12 the topside electron gas is not parasitic upon the runaway region that exists below 250 km. This is because the electron temperature (and hence the thermal conductivity) in this region is too low to permit an upward flux

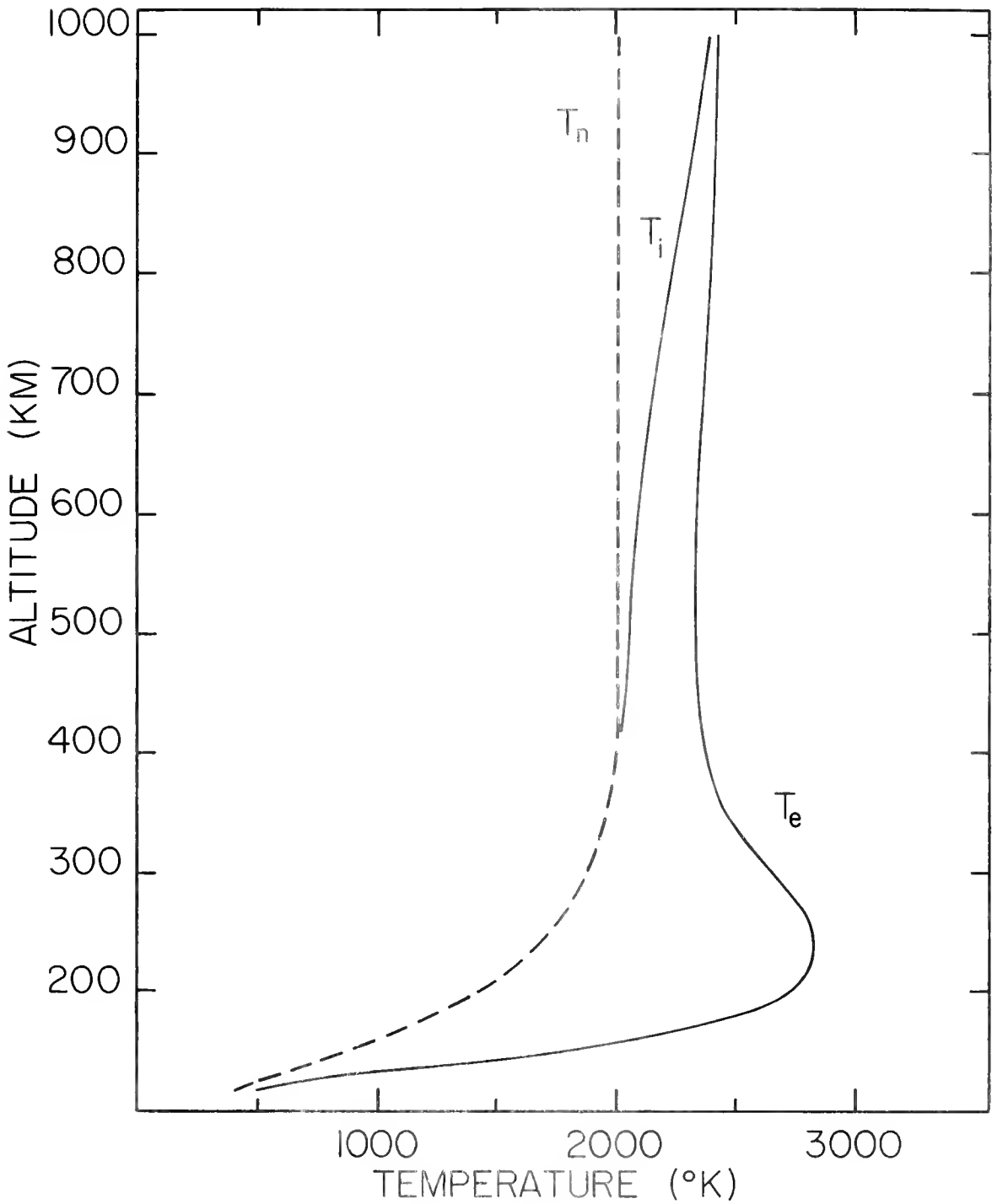

Figure 5.13 Ion and electron temperature profiles at sunspot maximum. Temperature of the neutral atmosphere is taken from Harris-Priester model S=250 at 1200 local time.

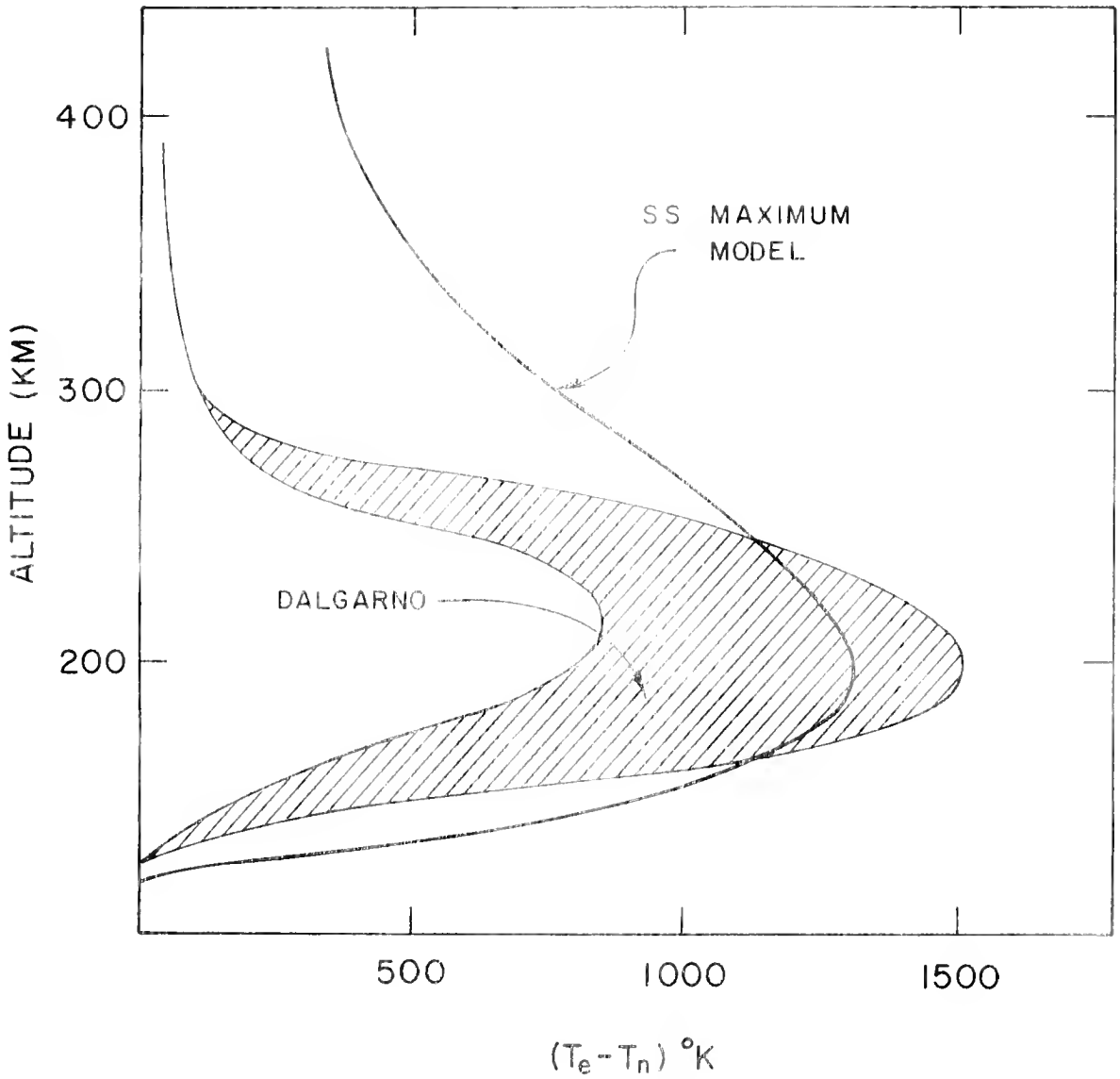

Figure 5.14 Comparison of the sunspot maximum electron temperature profile with those of Dalgarno et al. (1963).

of heat that is large compared to the capacity of the topside heat sink. For the sunspot maximum model, however, the electron temperature is sufficiently high, and parasitic overheating of the topside electron gas takes place. This is seen in Figure 5.15; a comparison of Figure 5.15 with Figure 5.12 shows that the parasitic coupling between the topside electron gas and that in the runaway region broadens the temperature maximum and prevents the re-establishment of thermal equilibrium at high altitudes.

5.8.3 Comparison With Observation

Figure 5.16 shows a comparison of the sunspot minimum model profiles with measurements of electron and ion temperatures made with a rocket-borne spherical ion trap launched from Eglin Air Force Base (Nagy, et al., 1963); excellent agreement is found. Recalling the fact that ion and neutral gas temperatures are equal in this region, it is possible that some of the discrepancy between predicted and measured ion temperature may be due to a neutral gas temperature that is slightly higher than that given by the model atmosphere, on that particular day.

A comparison with some profiles obtained from incoherent backscatter spectra (Evans, 1964) is given in Figure 5.17. In order to obtain the agreement shown here it was necessary only to scale down Q by a small factor. This is not unreasonable, since the measurements were made at a relatively high latitude station (Boston) over a 90 minute period centered at 0915 local time, with the resultant likelihood that there is some zenith angle attenuation of solar radiation. This measurement clearly illustrates the expected increase of ion temperature above 500 km due to the loss of thermal contact of the ions with the neutral atmosphere. Because of ground clutter, the measurements do not

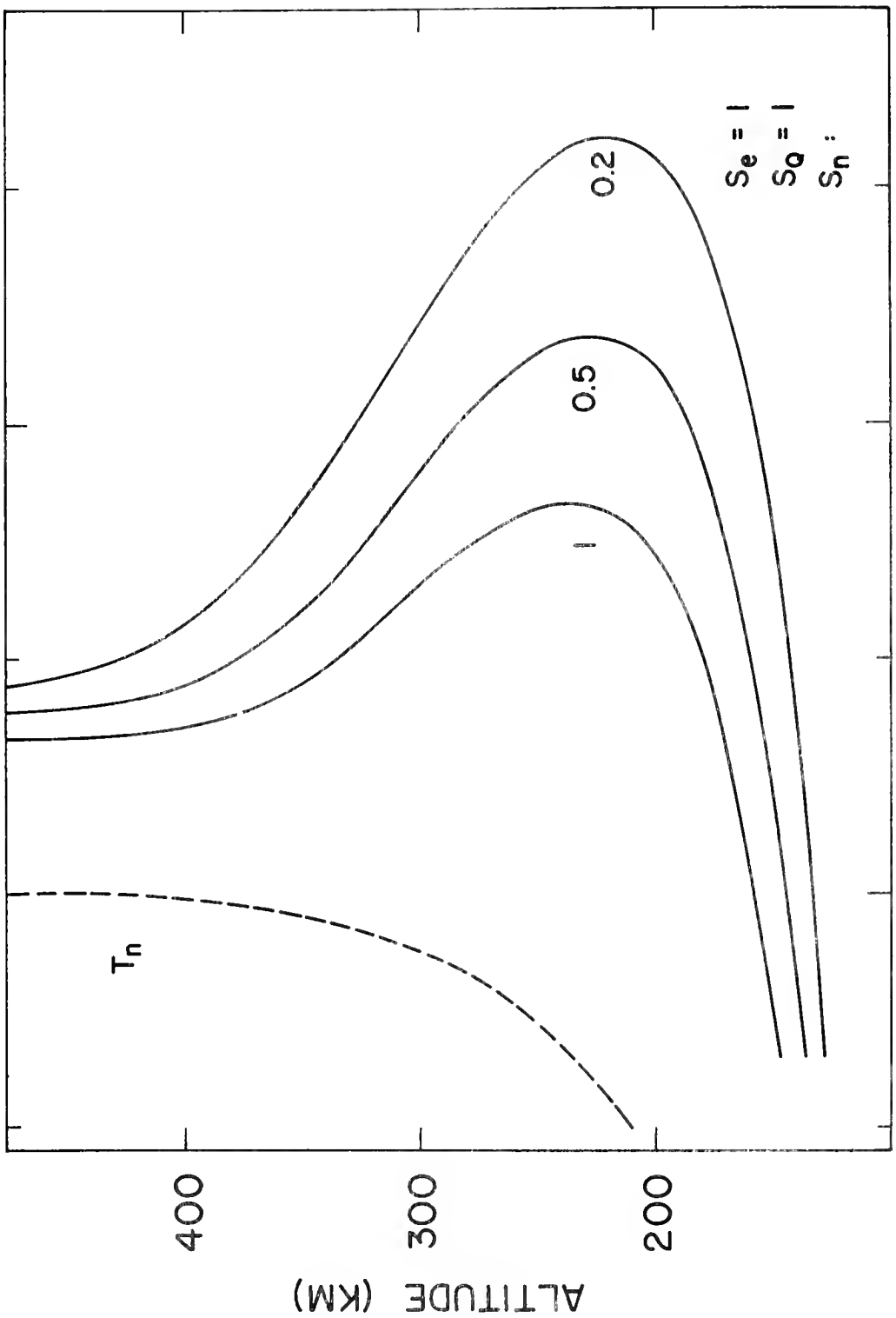

Figure 5.15 Illustrates parasitic overheating of the electron gas above 250 km in the sunspot maximum model. Compare with Figure 5.12.

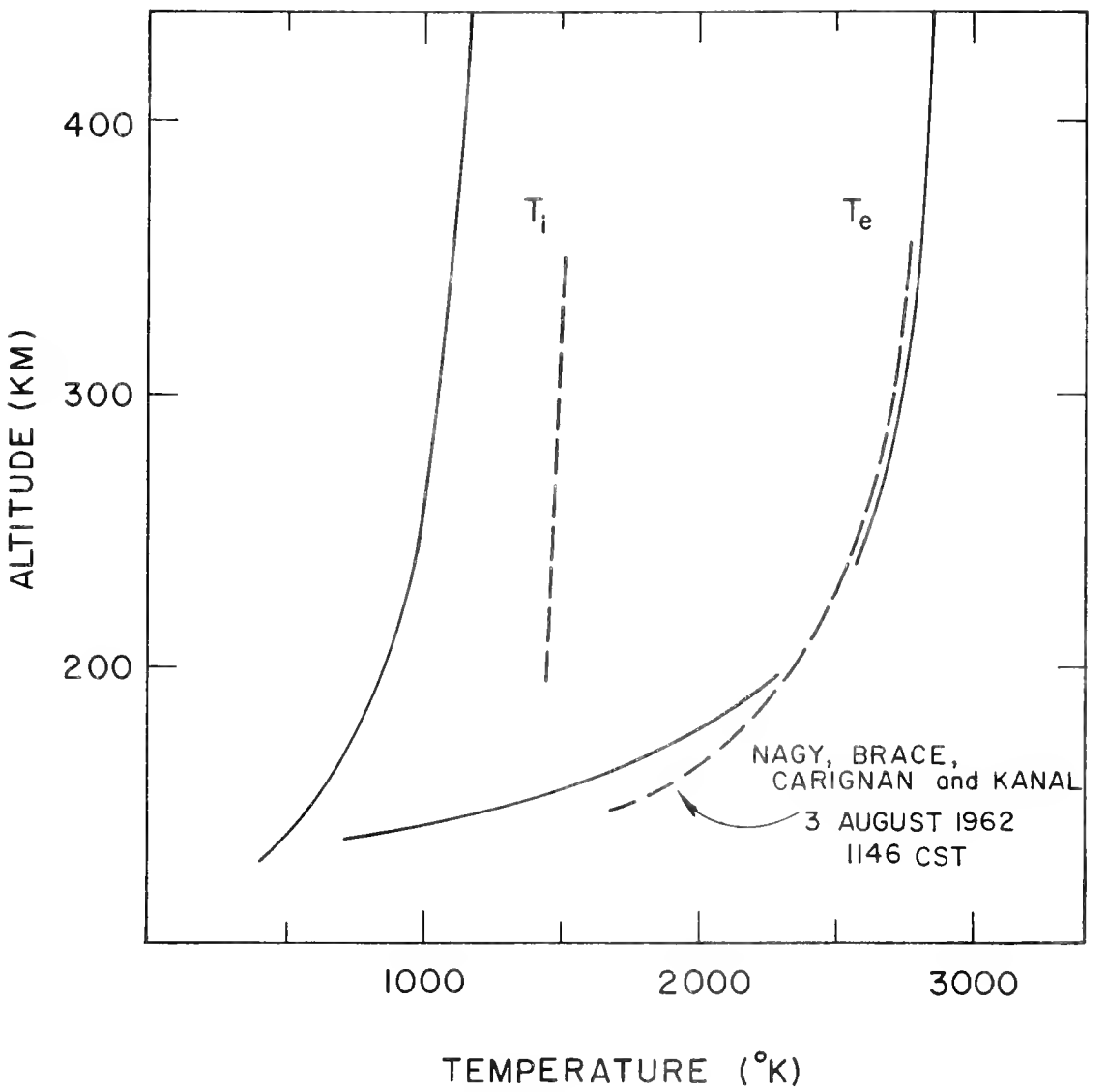

Figure 5.16 Comparison of the sunspot minimum temperature profiles (continuous line) with those obtained by Nagy et al. (1963).

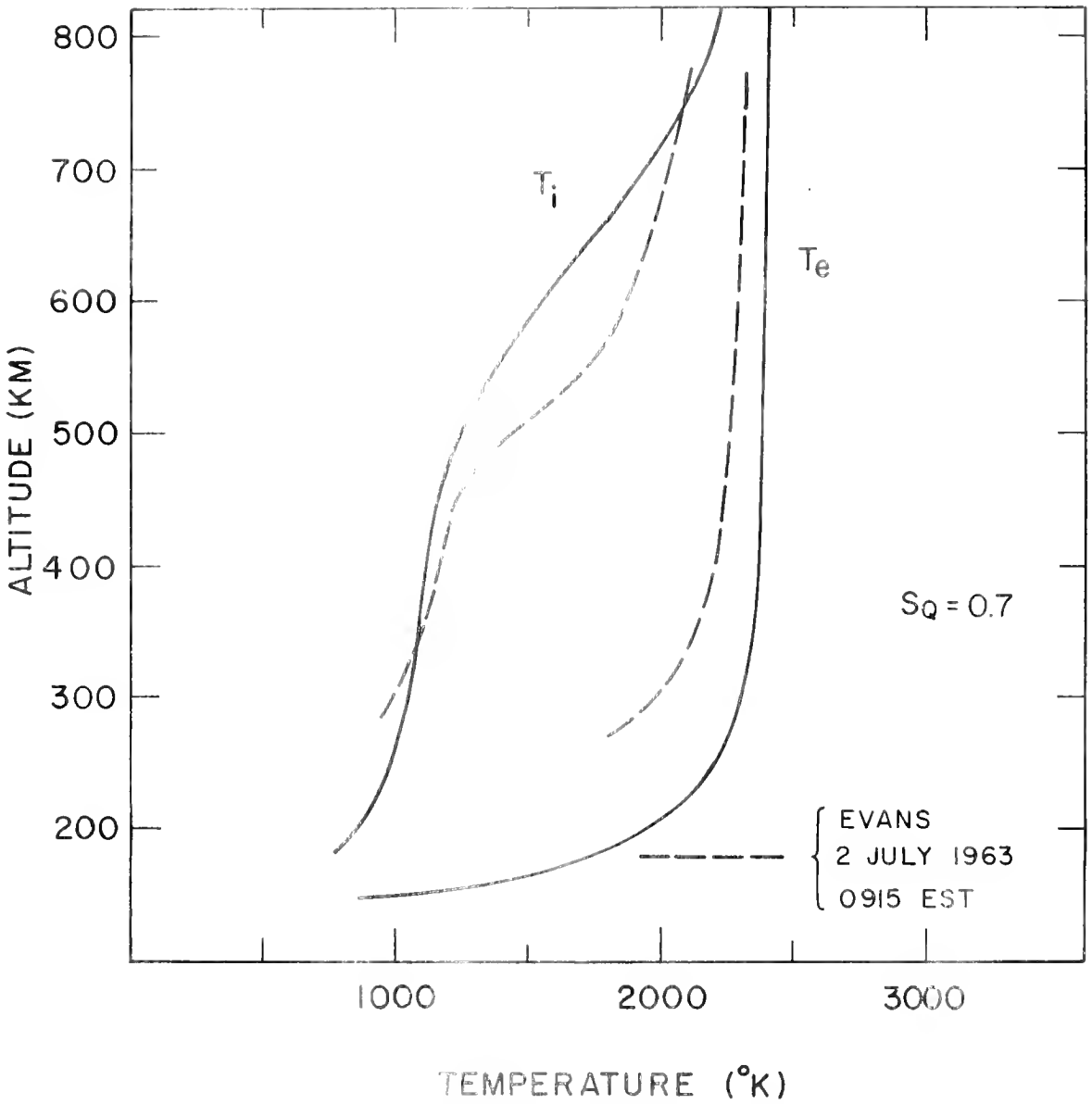

Figure 5.17 Broken lines represent ion and electron temperature profiles deduced from backscatter spectra (Evans, 1964). Theoretical profiles are those obtained by scaling the Q of the sunspot minimum model by a factor of 0.7.

extend below 250 km. However, there is evidence here that the departure from thermal equilibrium begins at a higher altitude than indicated by the theoretical profile. It is not possible to scale the parameters of the model in any reasonable way to duplicate this feature of the observed profile.

One of the notable exceptions referred to in Section 5.1--a mid-latitude measurement that does not show a monotonic increase of electron temperature with altitude--appears in Figure 5.18. This measurment was made with a dumbbell probe flown from Wallops Island (Brace, et al., 1962). The date of measurement is about midway between maximum and minimum of the solar cycle, and the heat input used in the theoretical model has been scaled accordingly. However, this modification of the sunspot minimum model was found to be insufficient, the resulting electron temperature profile still showing a monotonic increase with altitude. The added modification which produced the agreement shown in Figure 5.18 was the substitution for the model ionosphere of the sub-peak electron density profile measured at the time and location of the rocket launch. This seems to suggest that the relatively high electron density (1.6×10^6 cm^{-3} at the F2-layer maximum) present at the time was responsible for the existence of the sunspot maximum type electron temperature profile.

5.9 Thermal Capacity of the Ionosphere

The high electron and ion temperatures that are present in the upper atmosphere during the day have a profound influence on the structure of the ionosphere. They have virtually no effect on the structure and thermal balance of the neutral atmosphere, however. The reason for this is that the thermal capacity of the ionosphere is relatively small. The ion temperature is equal to that of the neutral atmosphere below about 500 km, so that most of the

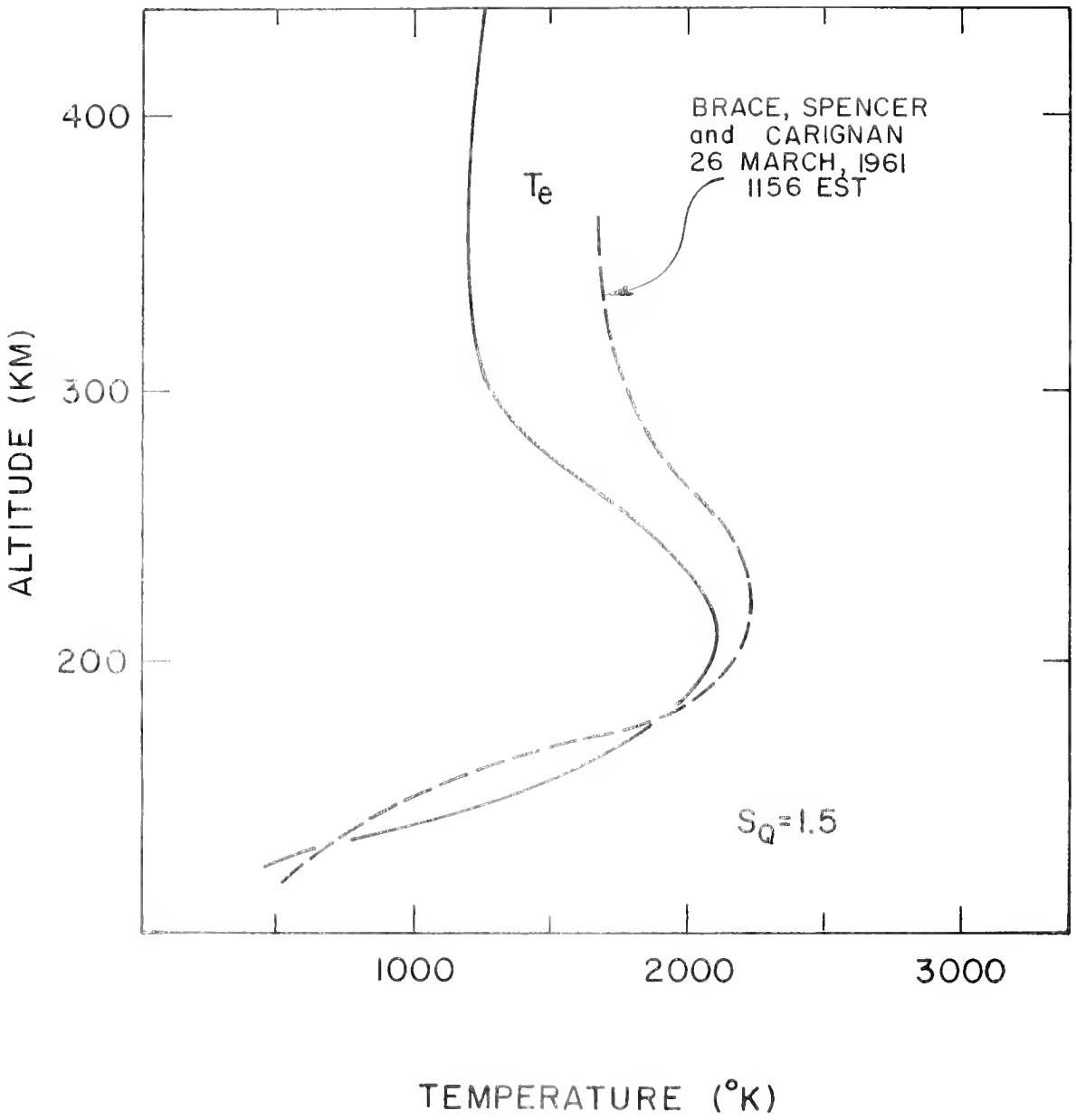

Figure 5.18 Electron temperature profile obtained by Brace et al. (1963) is represented by the broken line. The theoretical profile shown here was obtained by scaling the sunspot minimum Q by a factor of 1.5 and adopting for a sub-peak model ionosphere the electron densities measured at the time of the rocket launch.

thermal energy that could be transferred to the neutral atmosphere is stored in the electron gas. The energy thus stored in a unit column of the ionosphere is about 3/2 Nk $(T_e - T_n)$ where N is the total electron content of the column and $(T_e - T_n)$ is an average figure for the temperature difference. At sunspot minimum the daytime electron content is about 5×10^{12} cm^{-2}; a reasonable figure for $(T_e - T_n)$ inferred from the sunspot minimum model is $1500^{o}K$. For these values a thermal energy of 1.5 erg cm^{-2} is obtained. This is about equal to the energy conducted downward per second through the 150 km level by the neutral atmosphere (Hunt, 1962). At sunspot minimum, the total electron content is larger, but the difference $(T_e - T_n)$ is correspondingly smaller, so that the thermal energy stored in the electron gas remains about the same over the solar cycle. Thus, even in the event of a sudden collapse of electron temperature, the entire thermal energy content of the ionosphere could be disposed of in a few seconds without any effect on the structure of the neutral atmosphere.

The hypothetical sudden collapse just referred to raises the question of the magnitude of the time constant for cooling of the ionosphere. For the sunspot minimum model the integrated rate of cooling above 150 km is close to 1 erg cm^{-2} sec^{-1}. For a thermal content of 1.5 erg cm^{-2} a time constant of about a second is obtained. It is not significantly different at sunspot maximum. The rapid cooling rate indicated by the magnitude of this time constant makes it appear likely that thermal equilibrium will be established very shortly after sunset. This statement is of course based on the assumption that solar radiation is the only source of ionospheric heating. There is, however, another source of heat available to the nighttime ionosphere: the thermal energy stored in the protonosphere. This will be examined in the next chapter.

6. THE THERMAL STRUCTURE OF THE PROTONOSPHERE

6.1 Introduction

The region filled with a plasma of hydrogen ions and electrons that "floats" on top of the ionosphere was given the name protonosphere in Chapter 1. It was pointed out there that this ionization cannot move across lines of magnetic force and that for this reason it is preferable to examine the behavior of the ionization contained within a tube of magnetic force rather than that contained in a vertical column. Again in this chapter, the protonosphere will be thought of in terms of a protonospheric field tube that intersects unit area of the 1000 km surface at a geomagnetic latitude θ_o (see Figure 1.3).

The neutral atmosphere above 1000 km is very tenuous, and the photoionization rate is too low to serve as a heat source of any consequence. However, the solar radiation can supply heat to the protonosphere by means of the flux of photoelectrons that escape the F2-region of the ionosphere. These photoelectrons stream through the base of the field tube at the 1000 km surface and follow the contour of the tube up to the equatorial plane. The magnitude of the flux decreases with increasing distance above the 1000 km surface because photoelectrons are continually being reduced to thermal energies through coulomb interactions with the ambient electron gas. One objective of this chapter is to derive an expression for the rate at which the ionization is heated per unit length of the field tube by this process.

A further consequence of the low concentration of neutrals at these altitudes is that cooling of the electron gas cannot take place. This has two implications. The first is that ion and electron temperatures will be

equal in the protonosphere. Evidence for this has already been seen in the preceding chapter. The second implication is that a positive temperature gradient will be found in the protonosphere in order that the heat generated there may be conducted downward into the ionosphere where it can be transferred to the neutral atmosphere. An estimate of this heat flux is made in this chapter, and by integration of the heat conduction equation the temperature distribution within a mid-latitude protonospheric field tube is obtained.

Once the temperature distribution is known it is possible to estimate the thermal energy stored in a protonospheric field tube. This energy turns out to be of the order of the thermal energy stored in the ionosphere. At night this energy is released by thermal conduction downward into the ionosphere. Possible effects on the structure of the nighttime ionosphere are discussed.

6.2 Physical Assumptions

The problem of determining the rate of heat input to a given volume of a protonospheric field tube is solved in much the same way as the problem of heat input to unit volume of the ionosphere. The principal difference is that in the protonosphere it is necessary to work in terms of arc length along a field line rather than altitude. The expression for the distribution of ionization as a function of arc length along a field line is, however, sufficiently complicated that when integration over arc length is required it is necessary to resort to numerical methods. In view of the fact that a large number of integrations would be required in order to exhibit the dependence of the heating rate and the heat flux on the initial energy of the photoelectrons, this is to be avoided. The simplest assumption will be made,

namely, that the electron concentration is constant throughout the field tube. This assumption is not so drastic as it may seem, since the high temperature, the small mass of the hydrogen ion, and the decrease of the acceleration of gravity with altitude all conspire to produce a very large scale height of some 3500 km or more for the ionization.

Only a very small fraction of the photoelectrons which escape from the ionosphere are produced above 1000 km. Accordingly, it is quite reasonable to assume that the flux across the base of the protonospheric field tube is the same as that which escapes from the infinite model of the topside ionosphere. The assumption will be made, then, that the magnitude of the flux entering the base of the field tube is that derived in Section 5.3.2, that is, $Hq(z_e)/6$ cm^{-2} sec^{-1}, and that these photoelectrons have the pitch angle and energy distribution given by Equation (5.29).

Dependent upon the assumed initial energy of the photoelectrons produced in the F2-region, a certain fraction of the photoelectrons entering the base of the protonosphere cross the equatorial plane and heat the conjugate protonosphere. In order that this additional heat source may be taken into account, the calculations will be carried out for an equinox noon situation; that is, all physical quantities are assumed symmetric with respect to the geomagnetic equator.

6.3 Heating of the Protonosphere

6.3.1 The Heating Rate

For the photoelectron spiraling upward along a line of force in the protonosphere, Equation (5.3) takes the form

$$\frac{dW}{ds} = - \frac{Kn_e}{W \cos \alpha} \qquad (6.1)$$

where s is arc length along the line of force measured upward from the 1000 km surface. If the energy of the photoelectron is W_o at this surface, it will have at distance s an energy

$$W = W_o \left[1 - \left(\frac{E}{W_o} \right)^2 \frac{s}{R \cos \alpha} \right]^{1/2}, \qquad (6.2)$$

which follows from integration of Equation (6.1) under the assumption of constant electron density. In this expression the quantity

$$R = E^2/2 K n_e$$

is the range of a photoelectron with zero pitch angle and the maximum energy, E.

The number of photoelectrons crossing the unit area at the base of the tube per unit time and per unit interval of pitch angle and energy is given by the product of the number flux of escaping particles, $Hq(z_e)/6$, and the distribution function $\Psi(W_o \alpha)$ defined by Equation (5.29). The number per unit interval of pitch angle and energy which pass through the cross-sectional area of the tube per unit time at distance s is found by solving Equation (6.2) for W_o and substituting in Equation (5.29):

$$\frac{Hq(z_e)}{6}\Psi(W,\alpha) = 4Hq(z_e)g(\alpha)\cos^2\alpha \left[\frac{1}{E} \left(\frac{W^2}{E^2} + \frac{s}{R \cos \alpha} \right)^{1/2} \left\{ 1 - \left(\frac{W^2}{E^2} + \frac{s}{R \cos \alpha} \right) \right\} \right]. \qquad (6.3)$$

It can be seen from Equation (6.2) that there are now present no photoelectrons with pitch angles greater than a critical angle α^* given by

$$\cos \alpha^* = s/R,$$

and that, of the photoelectrons with pitch angle α, there are none present with energy that is less than a critical energy W* given by

$$W^* = E\left[1 - \frac{s}{R\cos\alpha}\right]^{1/2} .$$

The rate of heat input Q(s) to a volume of the field tube defined by the cross-sectional area of the tube and unit length along the direction of the field line is given by the product of $[Hq(z_e)/6]\,\Psi(W,\alpha)$ and $\left|dW/ds\right|$ integrated over energy and pitch angle:

$$Q(s) = \frac{2H}{E} Eq(z_e) \int_0^{\alpha^*} g(\alpha)\cos\alpha\, d\alpha \int_0^{W^*} \left\{1 - \left(\frac{W^2}{E^2} + \frac{s}{R\cos\alpha}\right)\right\} dW . \qquad (6.4)$$

Integration over energy gives

$$Q(s) = \frac{4H}{3R} Eq(z_e) \int_0^{\alpha^*} \left\{1 - \frac{s}{R\cos\alpha}\right\}^{3/2} \cos\alpha\, g(\alpha)\, d\alpha . \qquad (6.5)$$

Now, with the aid of the further substitutions:

$$L^2 = s/R,$$

$$\bar{\alpha} = \pi - \alpha,$$

and the assumption of an isotropic pitch angle distribution, Equation (6.5) becomes

$$Q(s) = \frac{2H}{3R} Eq(z_e) L^2 \int_{\bar{\alpha}^*}^{\pi} \left\{1 + L^2\sec\bar{\alpha}\right\}^{3/2} \frac{\cos\bar{\alpha}}{L^2} \sin\bar{\alpha}\, d\bar{\alpha} . \qquad (6.6)$$

The integral in Equation (6.6) is formally the same as that which appears in the equation for the non-local heating rate, Equation (5.25). Upon evaluation of this integral the expression for the heating rate is obtained:

$$Q(s) = \frac{H}{R}\, Eq(z_e)\, L^2 \left[\frac{2}{3}\,(1-L^2)^{1/2} \left\{ \frac{1}{4L^2} - \frac{5}{8} \right\} + \frac{L^2}{4} \log \left| \frac{1 + (1-L^2)^{1/2}}{L} \right| \right]. \quad (6.7)$$

Since Q(s) is the rate of heat input to a volume defined by the cross-sectional area of the field tube and unit length along the direction of ds, integration of Equation (6.7) over s from the base of the protonosphere to the equatorial plane will give the total rate of heat input to this half of the tube. There is also a contribution from photoelectrons which escape the conjugate ionosphere and cross the equatorial plane. For the assumed equinox noon condition, this contribution can be included by extending the integration across the equatorial plane down to the conjugate base of the field tube.

Thus, the protonospheric field tube intersecting unit area on the 1000 km surface at geomagnetic latitude θ_o and bounded by the equatorial plane is heated at a rate

$$\Phi(s) = \int_0^{2S} Q(s)\, ds, \quad (6.8)$$

where S is the distance measured along a field line from the base of the field tube to the equatorial plane. Equation (6.8) leads to integrals that are of much the same form as those encountered earlier, and the details will not be given here. When these integrals are evaluated, there results

$$\Phi(S) = HEq(z_e) \left[(1-L^2)^{1/2} \left\{ -\frac{4}{45} + \frac{13}{45} L^2 - \frac{11}{30} L^4 \right\} + \frac{L^6}{6} \log \left| \frac{1+(1-L^2)^{1/2}}{L} \right| + \frac{4}{45} \right], \quad (6.9)$$

where in this expression

$$L^2 = \frac{2S}{R} \quad . \tag{6.10}$$

It has been assumed here that the range (which is a function of the initial energy of the photoelectrons) is such that $R \geq 2S$. Note that as $R \rightarrow 2S$,

$$\Phi\,(S) \longrightarrow \frac{4}{45}\,\mathrm{HEq}(z_e),$$

that is, all of the escaping photoelectrons are stopped by the protonosphere before reaching the conjugate ionosphere.

6.3.2 The Heat Flux

For the assumption of negligible cooling to the neutral atmosphere above the 1000 km surface, Equation (6.9) represents the rate at which thermal energy is conducted downward through unit area of this surface by the electron gas. It is the purpose of this section to obtain an estimate of this quantity for the field tube with base at geomagnetic latitude 40° ($I \approx 70^{\circ}$).

It is quite sufficient for the present purpose to use the centered dipole approximation for the earth's magnetic field. The arc length S measured along a field line from the 1000 km surface at geomagnetic latitude 40° to the equatorial plane is about 8000 km. With $n_e = 3 \times 10^3$ cm^{-3} as the assumed constant number density of electrons within this field tube, the parameter L defined by Equation (6.10) is given by

$$L^2 \approx \frac{20}{E^2} \quad .$$

The factor enclosed within the square brackets in Equation (6.9) may be called

the transparency factor of the protonosphere. Figure 6.1a shows the transparency factor as a function of the photoelectron initial energy for this model.

The heat flux is given by Equation (6.9) summed over the spectrum of photoelectron initial energies. As pointed out previously, this is not well known, and the fact that the escape level is also a function of photoelectron initial energy introduces such practical difficulty that no attempt will be made to perform this summation. Instead, the heat flux will be estimated by a procedure consistent with that used to obtain the non-local heating rate in Chapter 5: it will be assumed that the escape level is at 300 km and that E is a mean energy defined by the ratio of heating rate to photoelectron production rate at this level. The rate of production of photoelectrons at 300 km deduced by Dalgarno et al. (1963) is about 2×10^2 cm^{-3} sec^{-1}, while the heating rate deduced for this altitude is about 1.5×10^3 ev cm^{-3} sec^{-1}. The mean initial energy is therefore approximately 10 ev. When this value of E is substituted into the transparency factor, the heat flux is given by

$$\Phi \approx HEq(z_e) \; [0.05]. \tag{6.11}$$

For sunspot minimum the production rate just quoted should be scaled by a factor of 0.7, and H ≈ 60 km. Upon substitution of these quantities into Equation (6.11), a figure of 3×10^8 ev cm^{-2} sec^{-1} is obtained for the rate at which heat is conducted downward through unit area of the 1000 km surface at geomagnetic latitude 40^o during the day.

Some assessment of the accuracy of this result can be made at this time. Although the transparency factor is a sensitive function of photoelectron

initial energy, this is not the case with the heat flux. This is seen in Figure 6.1b. The result just obtained would not therefore appear to depend strongly on the assumed mean initial energy. There is, however, another factor to be considered: the relation of the initial energy to the escape level. When 10 ev is substituted for the initial energy in Equation (5.7), it is found that $n_e(z_e) \approx 2 \times 10^6$ cm^{-3}. This figure is reasonably consistent with the observed electron density at 300 km for daytime sunspot maximum conditions. At sunspot minimum, however, this is a factor of 5 greater than the density at the F2-layer maximum. What this means, in effect, is that the total electron content of the ionosphere at sunspot minimum is insufficient to stop a photoelectron of zero pitch angle and 10 ev initial energy. The situation, therefore, is one in which the escape flux is limited by the neutral atmosphere rather than the ionosphere. This opens the possibility that the escape flux may be far in excess of the $Hq(z_e)/6$ cm^{-2}sec^{-1} obtained when it is assumed that $z_e = 300$ km. This increase would affect the factor appearing in front of the square brackets in Equation (6.11) with the result that the heat flux estimate of 3×10^8 ev cm^{-2} sec^{-1} may be seriously in error for sunspot minimum.

This possibility will be examined in the next section by deriving an estimate of the rate at which photoelectrons escape from a model atmosphere consisting only of neutral constituents.

6.3.3 Effect of the Neutral Atmosphere

The question to be answered here is whether or not the neutral atmosphere is sufficiently transparent to a flux of photoelectrons that those photoelectrons which originate below the 300 km level can make an appreciable contribution to

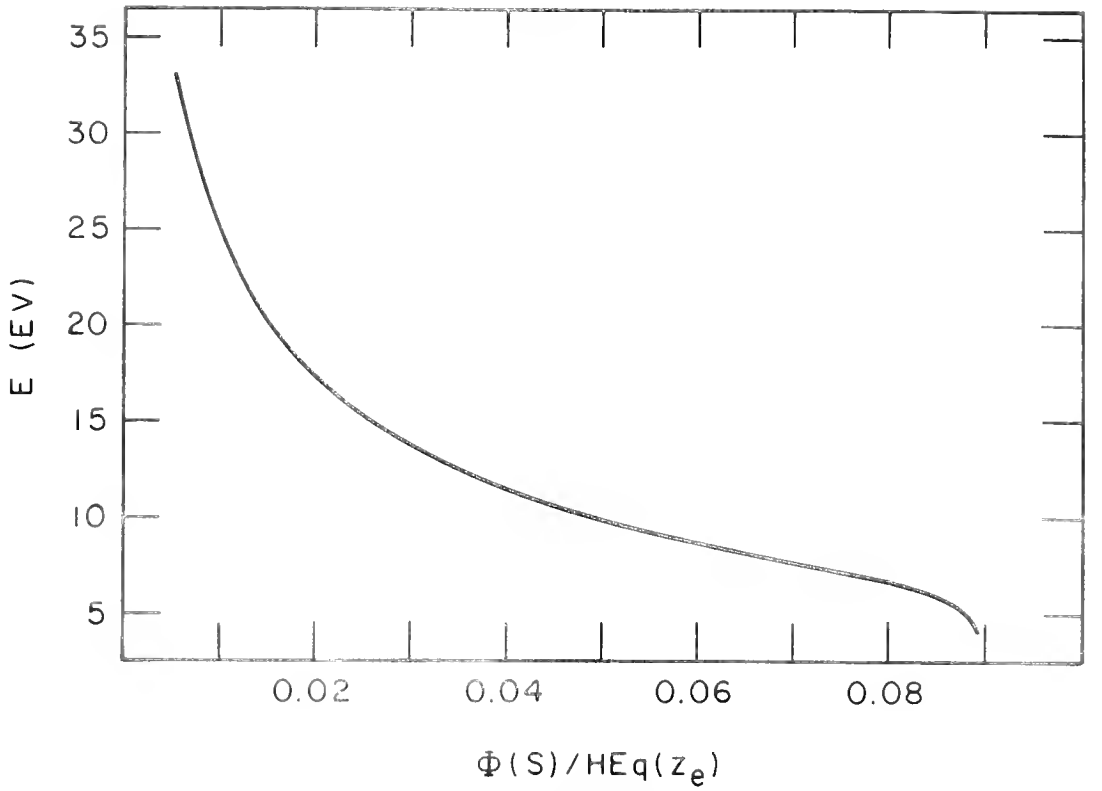

Figure 6.1a. The transparency factor as a function of photoelectron initial energy.

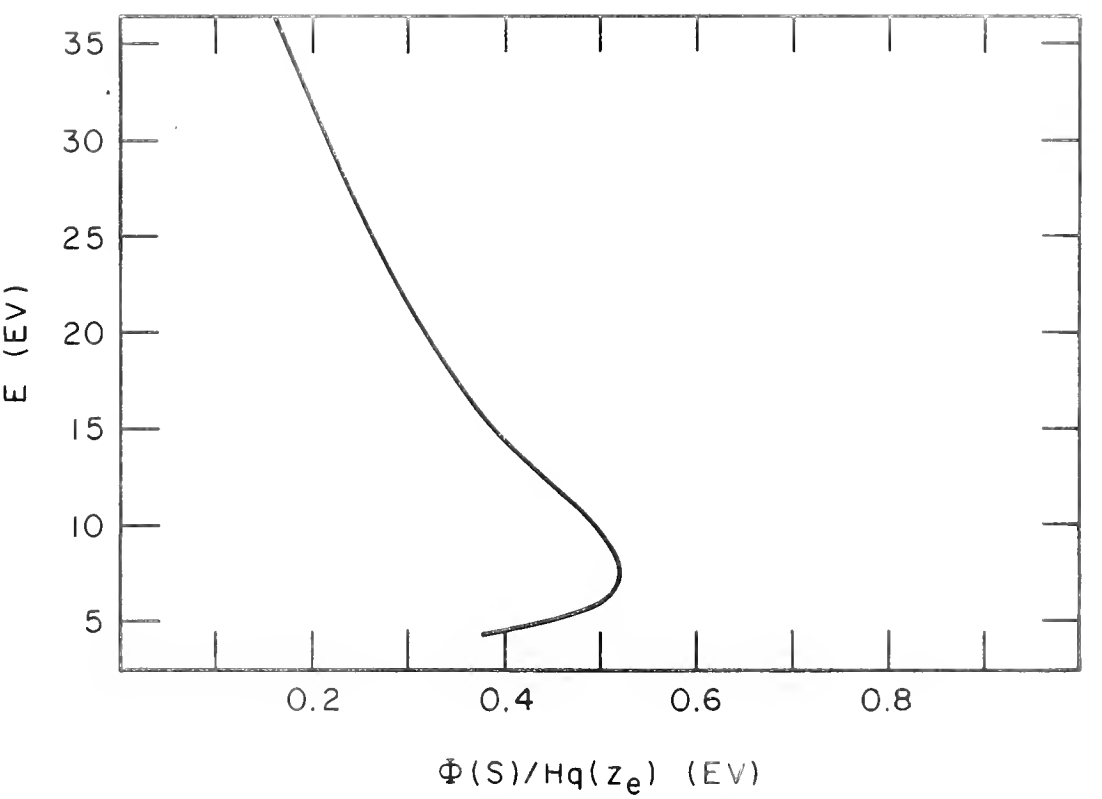

Figure 6.1b. Dependence of heat flux at the base of the tube on photoelectron initial energy.

the escape flux. It is therefore necessary to take into account the dependence of the production function upon altitude in the region below the 300 km level. For this purpose the full "Chapman function" for the rate of production of ionization will be used (Chapman, 1931):

$$q = q_m \exp\left[1 - (z-z_m)/H - \exp\left\{-(z-z_m)/H\right\}\right], \qquad (6.12)$$

where q_m is the maximum rate of production, and H is the scale height of the ionizable constituent.

We consider again the simple electron beam approximation and neglect the effect of the magnetic field. The equation of continuity for the flux G of photoelectrons in the upward direction is

$$\frac{dG}{dz} = \frac{q}{2} - \sigma n G, \qquad (6.13)$$

where n is the number density of the neutral atmosphere. The second term on the right-hand side of this equation follows from the assumption that photoelectrons are scattered out of the beam at a rate per unit length that is jointly proportional to the magnitude of the flux and the concentration of the scattering agent with constant of proportionality σ.

The further assumption will now be made that the altitude distribution of the neutral atmosphere is exponential with scale height equal to that of the constituent providing the photoelectrons. In this case, Equation (6.13) may be integrated by means of an integrating factor to obtain

$$G(z) = C + \frac{Hq_m}{2(1+\sigma n_m H)} \exp\left[1 - \exp\left\{-(z-z_m)/H\right\}\right], \qquad (6.14)$$

where C is a constant of integration. This constant must be zero in order to satisfy the lower boundary condition: $G \to 0$ as $z \to -\infty$. The expression for the escape flux is obtained by taking the limit of Equation (6.14) as $z \to \infty$:

$$G(\infty) = \frac{H q_m e}{2(1 + \sigma n_m H)} \quad . \tag{6.15}$$

The proportionality constant σ in the collision term of Equation (6.13) is commonly expressed in terms of the probability of collision P_c (see e.g. Brown, 1959):

$$\sigma n = p \left(\frac{273}{T}\right) P_c \quad , \tag{6.16}$$

where p and T are the pressure (in millimeters) and temperature of the scattering gas. Measurements of P_c have been obtained for an electron beam in a gas of O_2 and in a gas of N_2. If the narrow resonance for 3 ev electrons in N_2 is ignored, P_c is very nearly independent of energy and has practically the same value for both N_2 and O_2; the value $P_c = 35$ has been chosen from inspection of Figure 1.11 of Brown (1959).

The production maximum is located near the 180 km level. From the Harris-Priester model $S = 100$ at local noon, $p \approx 1.6 \times 10^{-6}$ mm, $T \approx 800^{\circ}K$, $n = n(O_2) + n(N_2) \approx 10^{10}$ cm^{-3}, at this altitude. The average scale height of the atmosphere above 180 km at this time is close to 50 km. Substitution of these data and the value of P_c into Equation (6.16) gives $\sigma n_m H \approx 90$. Further substitution of this value into Equation (6.15) gives for the escape flux

$$G(\infty) \approx \frac{1}{90} \left\{\frac{H q_m \epsilon}{2}\right\} \quad .$$

Experimental values of P_c are determined from measurement of the change in current when an electron beam passes through different distances of the scattering gas. Only one collision is sufficient to scatter an electron out of this beam. In the atmosphere, however, a single collision does not sufficiently alter the vertical velocity component of an electron initially moving straight upward to preclude the possibility that this electron will escape. The use of P_c in Equation (6.15) therefore leads to an underestimate of the escape flux. Following the suggestion of Hanson (1963), we shall assume that a photoelectron is not contained by the neutral atmosphere unless three such collisions are made. The derived escape flux should therefore be increased by a factor of 3.

When this correction has been included, the estimate of the escape flux from a model atmosphere consisting only of neutral constituents and a Chapman production function is given by

$$G(\infty) \approx \frac{1}{10}\left\{\frac{Hq_m}{2}\right\}. \qquad (6.17)$$

Now with $z_e = 300$ km and the assumed average scale height of 50 km for the region between 180 km and 300 km it follows that $q(z_e) \approx 10^{-1} q_m$. Reference to Equation (6.17) then shows that the escape flux is close to $Hq(z_e)/2$ cm^{-2} sec^{-1} if z_e is chosen to be 300 km. The quantity $Hq(z_e)/2$ is, however, just the escape flux deduced for the electron beam approximation in a model ionosphere of scale height 2H.

It follows, then, that the escape flux at sunspot minimum is correctly given by $Hq(z_e)/6$ if z_e is chosen to be 300 km, even though the electron

content is so low at this time that escape is limited by the neutral atmosphere and not by the ionosphere. Thus, the estimate of the heat flux through the 1000 km surface derived in the preceding section remains unchanged and stands at 3×10^8 ev cm^{-2} sec^{-1} at geomagnetic latitude 40° for daytime sunspot minimum conditions.

At sunspot maximum the electron content is sufficient that the escape of photoelectrons is limited by the ionosphere. The escape flux is still given by Equation (6.11); however, the scale height H is greater by a factor of 1.7 than at sunspot minimum, and the production rate must be scaled by the ratio of the intensity of the solar radiation at the 10.7 cm wavelength. These factors combine to increase the escape flux by a factor of 5 and lead to an estimate of 1.5×10^9 ev cm^{-2} sec^{-1} for the heat flux at sunspot maximum.

6.4 Temperature of the Protonosphere

The analysis of the heat balance of the ionosphere at sunspot minimum carried out in Chapter 5 resulted in a figure of about $2800^\circ K$ for the common electron-ion temperature on the 1000 km surface at middle geomagnetic latitudes. This figure was obtained for the case of no heat flux downward through the 1000 km surface. However, as stated earlier, the temperature of the topside ionosphere is relatively insensitive to a flux of the order of 3×10^8 ev cm^{-2} sec^{-1}, and the figure of $2800^\circ K$ will be taken as the temperature at the base of the protonosphere.

The temperature distribution along the magnetic field lines is set by the requirement that the heat supplied to a protonospheric field tube be conducted downward through the base of the tube. An upper limit to the temperature at the equatorial face of the tube with base at 40° geomagnetic latitude may be

obtained by integration of the heat conduction equation for an assumed flux of 3×10^{8} ev cm^{-2} sec^{-1} at all points along the field line located at the center of the tube. Integration of this equation is straightforward in this particular case and leads to the expression

$$T_e(S) = \left[T_o^{7/2} + (1.54 \times 10^{3})\, S \right]^{2/7}, \tag{6.18}$$

where T_o is the temperature at the base and S is the length along the line of force from the base to the equatorial plane. Substitution of $T_o = 2800^{\circ}K$ and $S = 8 \times 10^{8}$ cm into Equation (6.18) gives $T_e(S) \approx 3500^{\circ}K$. Thus, a temperature difference of at most $700^{\circ}K$ is required between the two ends of the protonospheric field tube with base at 40° geomagnetic latitude in order that the heat flux through the base just equals the rate of heat input to the tube.

The figure for the temperature difference just obtained is an overestimate, since the flux decreases with length along the tube. To obtain the correct figure it is necessary to integrate the heat conduction equation using the expression for the flux as a function of arc length. In this case Equation (6.18) takes the form

$$T_e(S) = \left[T_o^{7/2} + (5.12 \times 10^{-6}) \int_0^S \frac{\Phi(s)}{A(s)}\, ds \right]^{2/7}. \tag{6.19}$$

The factor A(s) is the cross-sectional area of the tube and appears in the integrand because the heat conduction equation is a relation between the temperature gradient and the rate of transport of thermal energy per unit area.

The expression for $\Phi(s)$ follows from integration of $Q(s)$ over the portion of the field tube above the point distance s from the base:

$$\Phi(s) = \int_{s}^{S} Q(s)\, ds + \int_{S}^{2S-s} Q(s)\, ds \quad . \tag{6.20}$$

The second integral in this equation represents the contribution from heating by photoelectrons originating in the conjugate hemisphere. For s = o Equation (6.20) reduces to Equation (6.8). Equation (6.20) may be put into the form

$$\Phi(s) = \int_{o}^{2S-s} Q(s)\, ds - \int_{o}^{s} Q(s)\, ds \quad , \tag{6.21}$$

and the right-hand side may be evaluated from Equation (6.9) if the parameter L^2 is taken to be (2S - s)/R for the first integral and s/R for the second. Figure 6.2 shows the behavior of $\Phi(s)$ for S = 8000 km.

The cross-sectional area of the tube cannot be expressed as an analytic function of arc length along the field line. It is therefore necessary to tabulate A(s) as a function of s, and this can be done in the following manner:

By definition, a field tube always contains the same quantity of magnetic flux. Thus, if A is the cross-sectional area of the tube with unit area at r_o, θ_o, then

$$B(r,\theta)\, A\, (r,\theta) = B(r_o,\theta_o) \quad , \tag{6.22}$$

where B is the magnetic field, r the geocentric distance and θ the geomagnetic

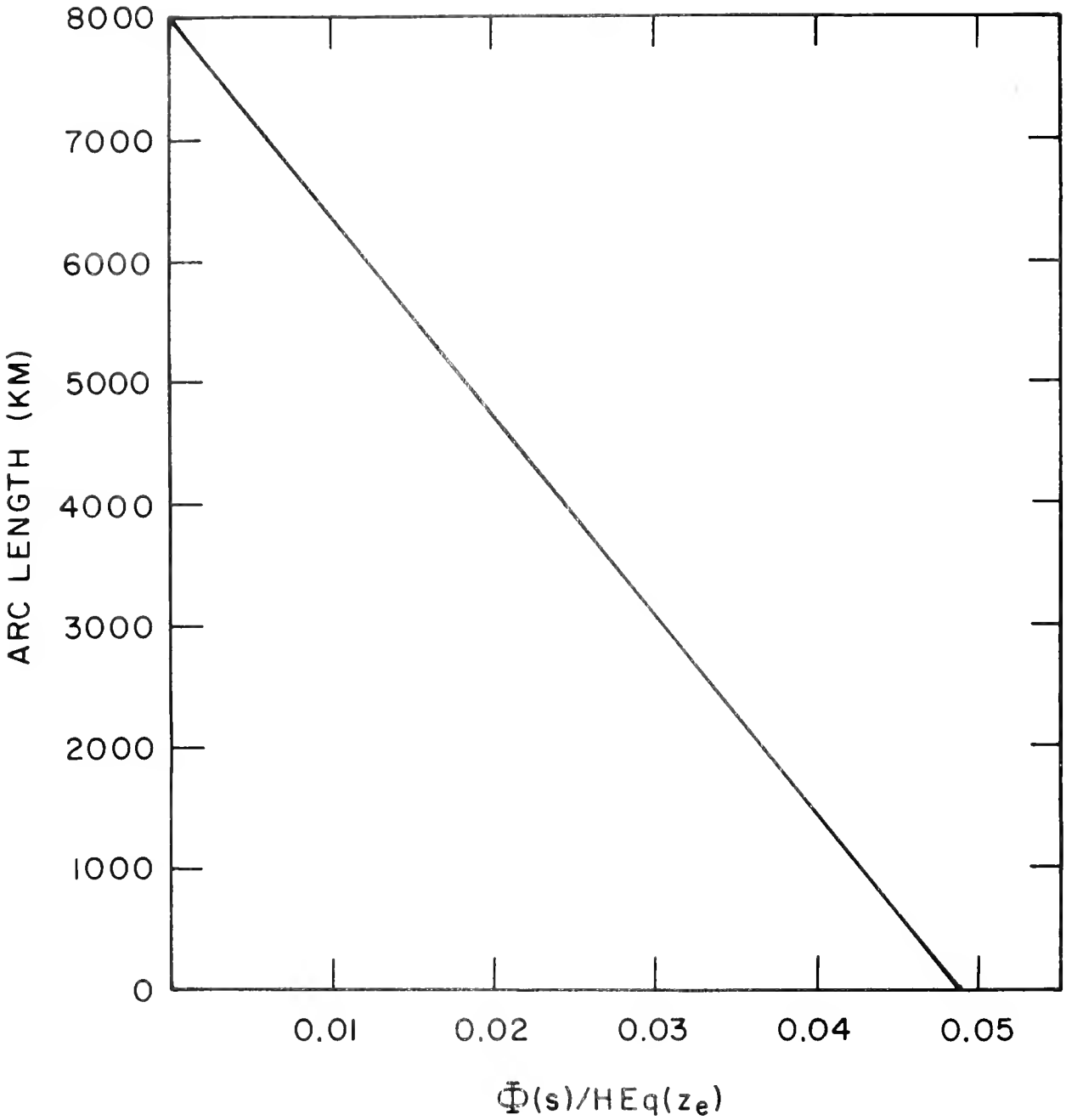

Figure 6.2 Rate of heat transport along the field tube.

latitude. For a dipole field

$$B(r,\theta) \propto \frac{(1 + \sin^2\theta)^{1/2}}{r^3} . \tag{6.23}$$

The equation of a line of force of the dipole field is

$$r = r_e \cos^2 \theta , \tag{6.24}$$

where r_e is geocentric distance to the point of intersection of the field line with the geomagnetic equator. This equation may be used to eliminate θ from Equation (6.23) to obtain

$$B(r) \propto \frac{\left\{4-3(r/r_e)\right\}^{1/2}}{r^3} , \tag{6.25}$$

and Equation (6.22) then gives

$$A(r) = \left(\frac{r}{r_o}\right)^3 \left[\frac{4-3(r_o/r_e)}{4-3(r/r_e)}\right]^{1/2} , \tag{6.26}$$

from which A may be tabulated as a function of r.

The expression for arc length along a line of force in the dipole field is

$$S(r_o,\theta_o) - s(r,\theta) = \frac{r_o}{2\sqrt{3}\cos^2\theta_o}\left[x + \sinh x \cosh x\right] , \tag{6.27}$$

where

$$\sinh x = \sqrt{3} \sin \theta$$

The variable θ can again be eliminated by substitution from Equation (6.24), and arc length may then be tabulated as a function of r.

The tabulated A(r) and s(r) are then combined to give A(s) in tabular form. The behavior of this function is illustrated in Figure 6.3. Comparison with Figure 6.2 will reveal that the divergence of the field tube and the decrease of $\Phi(s)$ are such as to reduce the heat flux by an order of magnitude over the base value only halfway up the tube.

The functions $\Phi(s)$ and A(s) so obtained may now be used to evaluate the integral that appears in Equation (6.19). This equation gives, in addition to the temperature in the equatorial plane, the temperature as a function of arc length. The function $I_e(s)$ is shown in Figure 6.4. It can be seen that most of the temperature change takes place in the lower one-third of the field tube. The temperature at the equatorial face of the tube is about $3000^{\circ}K$. Thus, when the divergence of the field tube and the attenuation of the flux of photoelectrons heating the protonosphere are allowed for, a temperature difference of only $200^{\circ}K$ is found between the base and the equatorial face of the tube with base at 40° geomagnetic latitude.

In view of the sensitivity of the 1000 km temperature to the assumed heat input function in the F region (see Figure 5.10), it is in the interest of completeness to consider the effect on the equatorial temperature of a change in temperature at the base of the tube. Equation (6.19) has been solved for various values of T_o within the range considered possible for the daytime 1000 km temperature at sunspot minimum. In each case, the flux has been scaled by the ratio of the assumed 1000 km temperature to $2800^{\circ}K$. This follows from the linear relation between this temperature and S_Q (see Figure 5.10) and the

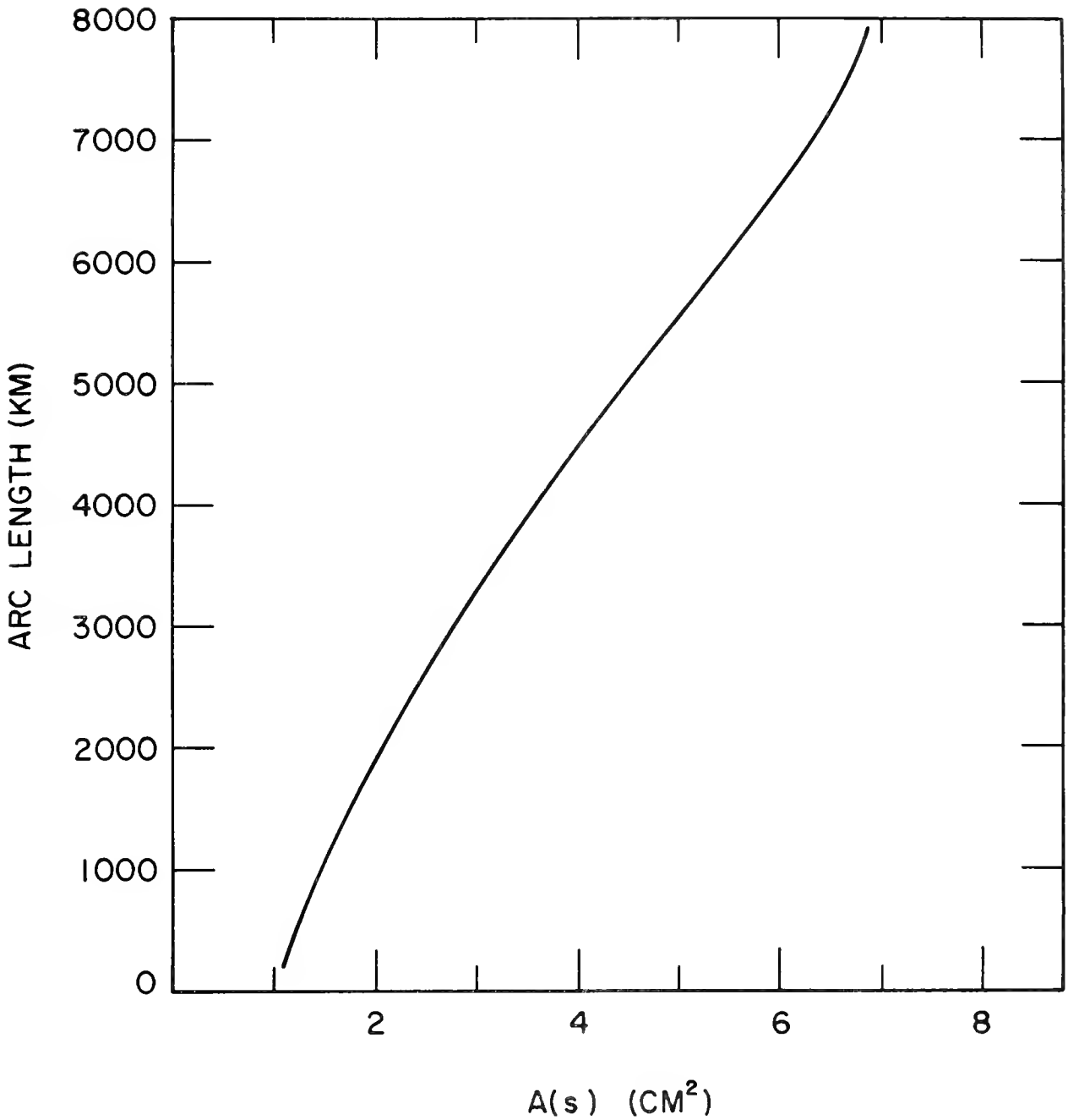

Figure 6.3 Cross-sectional area of the field tube with unit area at the base.

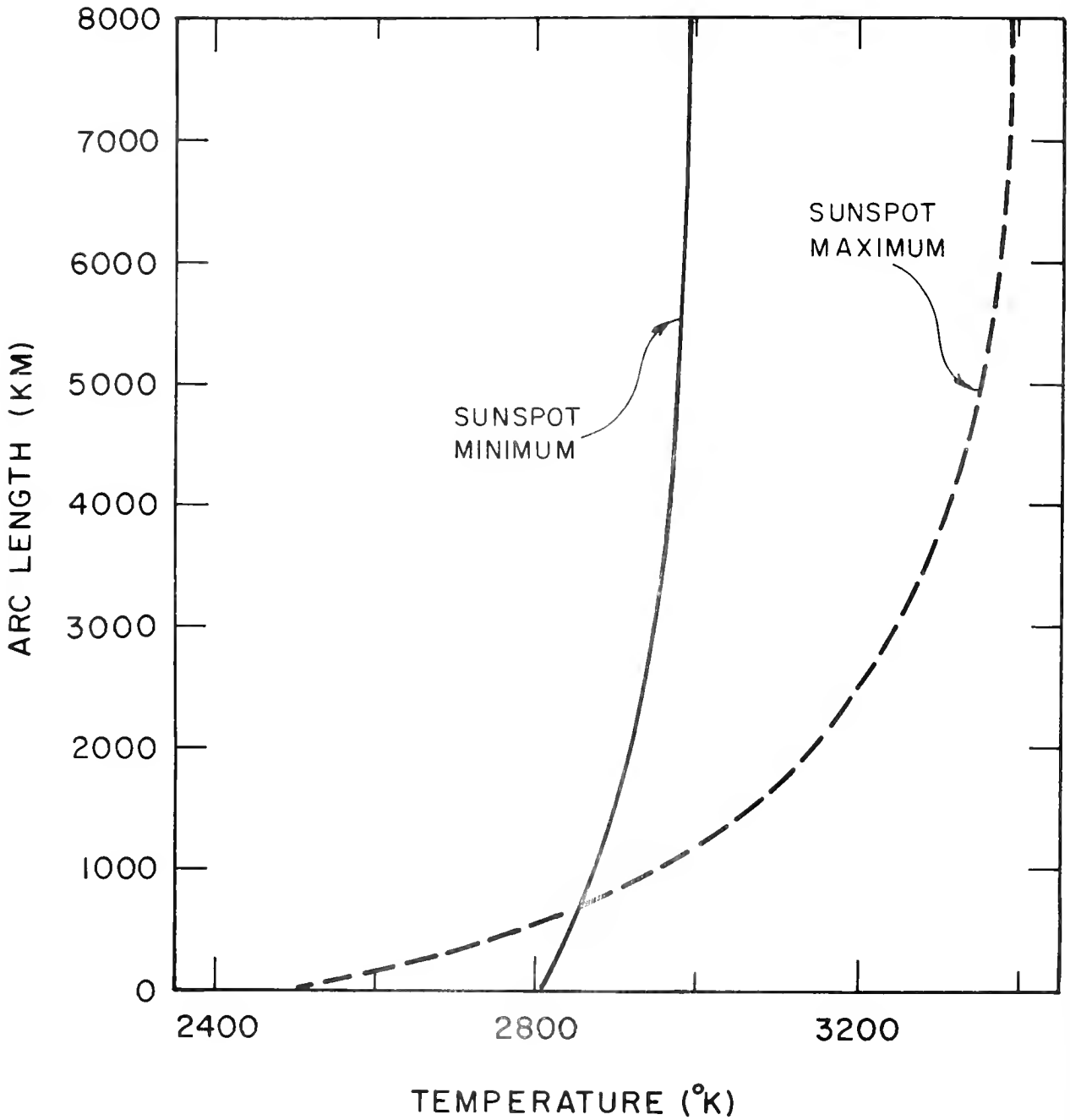

Figure 6.4 Temperature distribution along the field tube.

proportionality of $\Phi(s)$ and $q(z_e)$. The result of these calculations appears in Figure 6.5. A change of temperature at the base of the tube results in a change of temperature of nearly the same amount at all points along the tube. It may therefore be concluded that the linear relation between the 1000 km temperature and the scale factor for the F region heating rate deduced for the sunspot minimum model of Chapter 5 applies also to the temperature throughout the protonosphere. This behavior of the temperature profile along the field tube is a consequence of the combined effects of the high temperature dependence of the thermal conductivity and the relatively small heat flux. If the right-hand side of Equation (6.19) is expanded by the binomial theorem, there results

$$\frac{T_e(S)}{T_o} \approx 1 + \frac{2}{7}\left\{\frac{(5.12 \times 10^{-6})}{T_o^{7/2}} \int_o^S \frac{\Phi(s)}{A(s)}\, ds\right\}, \qquad (6.28)$$

Now for $T_o = 2200^oK$ the term enclosed by the curly brackets has the value 4.67×10^{-1}, while for $T_o = 3000^oK$ it has the value 2.17×10^{-1}. These values are proportional to the magnitude of the flux. Multiplication by the factor 2/7, which arises from the exponent of the temperature in the expression for the conductivity, reduces the second term of Equation (6.28) to a quantity that varies between about 10^{-1} and 6×10^{-2} over the range of T_o considered. Thus, although $T_e(S)/T_o$ increases with decreasing T_o, this ratio is still close to unity for the lowest value of T_o considered here.

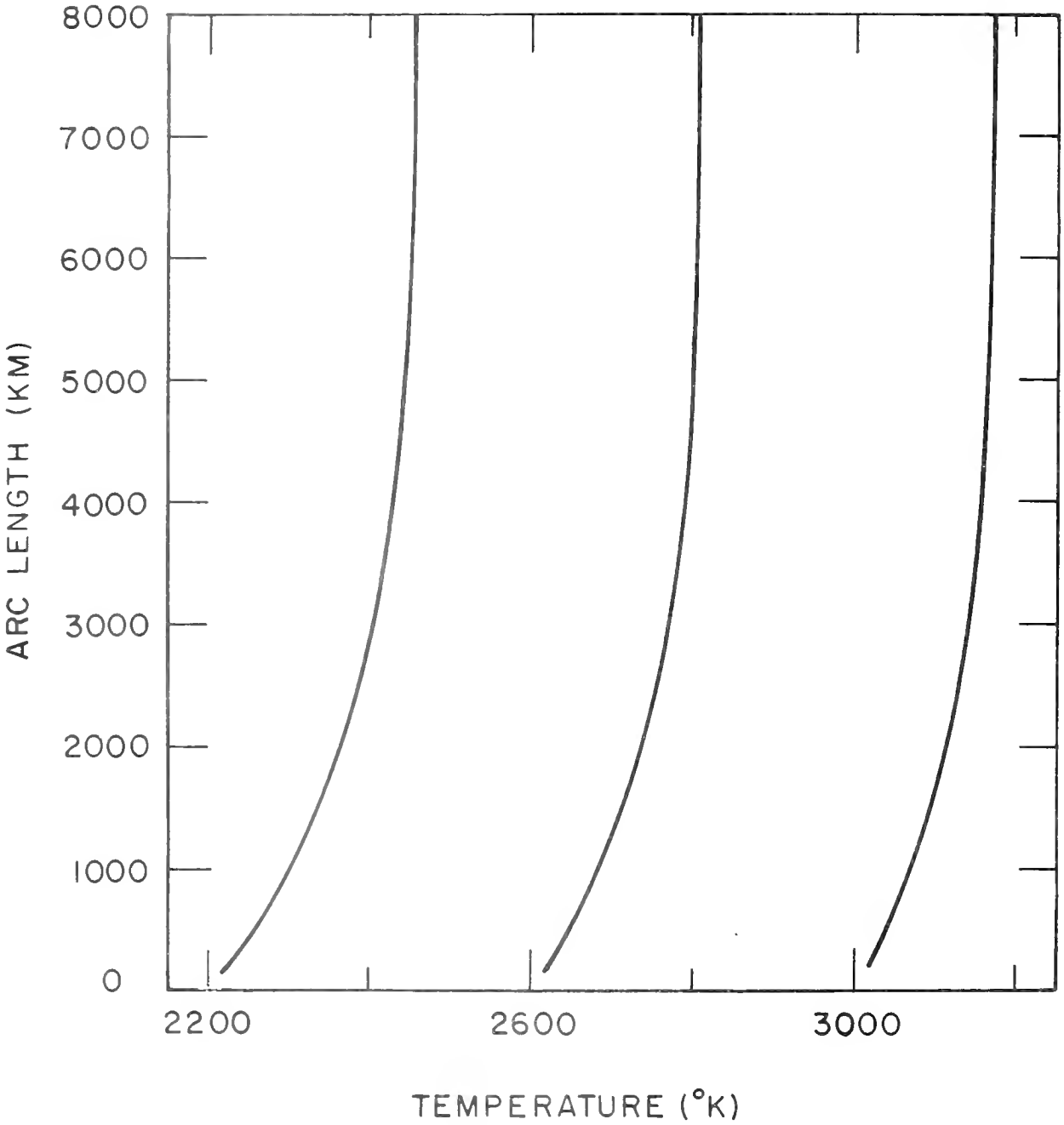

Figure 6.5 Temperature distributions along the field tube for sunspot minimum conditions, obtained by scaling Q.

Before this section is closed, some indication of the change in protonospheric temperature to be expected in the transition to sunspot maximum conditions should be included. The results of Chapter 5 indicate a temperature of about 2500°K at the base of the protonosphere at sunspot maximum. In Section 6.3, the heat flux through the base was estimated at 1.5×10^{9} ev cm^{-2} sec^{-1}. Since the temperature at the base was calculated under the assumption of no heat flux, the first question to answer is whether or not the estimated flux changes this temperature. To provide this answer, further integrations of the heat conduction equation were made for sunspot maximum conditions with this heat flux incorporated. No appreciable change in the 1000 km temperature was found. The principal reason for this is that the increase in electron content of the ionosphere at sunspot maximum (and the attendant absence of the runaway condition above 250 km) provides a heat sink with a capacity considerably in excess of this heat flux.

The temperature distribution that results from integration of Equation (6.19) when $T_o = 2500°K$ and a flux through the base of 1.5×10^{9} ev cm^{-2} sec^{-1} is assumed is shown by the broken line in Figure 6.4. The equatorial temperature is about 3400°K. No attempt has been made to take into account the change in electron content of the field tube over the solar cycle, as no observations are available on which to base an estimate of this change. A significant increase in the content should occur between sunspot minimum and sunspot maximum as a result of the increase in electron density at the F2-layer maximum over this period. The protonospheric temperature distribution that has been deduced may not be far in error, however, as the figure of 10^{3} cm^{-3} for the density in the equatorial plane (upon which the assumed mean density

f 3 x 10^3 cm^{-3} is based) is deduced from whistler observations made at a time hat is about midway between the maximum and minimum of the solar cycle.

.5 Nighttime Thermal Coupling Between the Ionosphere and the Protonosphere

It is well to begin this section with a review of the daytime situation for unspot minimum conditions. The daytime ionosphere is heated by the energy xtracted from photoelectrons released by the solar ultraviolet radiation and ools through collisional excitation of low energy molecular states by the elec- rons and through elastic collisions between ions and atomic constituents. Heat ransport by thermal conduction is an important process in maintaining a balance etween the rate of heating and the rate of cooling, and the electron tempera- ure above about 200 km is set by the requirement that most of the heat generated bove this level be conducted downward to the region between 120 and 140 km.

The mean free path for energy loss of a photoelectron becomes very large at ltitudes above that of the F2-layer maximum with the result that an upward flux f some 1 x 10^8 photoelectrons crosses unit area of the 1000 km surface. These hotoelectrons spiral upward along the lines of force of the geomagnetic field nd a large percentage are reduced to thermal energies before reaching the equa- orial plane. The heat thus produced in a protonospheric field tube cannot be ransferred to the neutral constituents, which are too few in number at these ltitudes, and must instead be conducted downward along the field lines through he 1000 km surface. When account is taken of the photoelectrons which cross he equatorial plane from the conjugate hemisphere, a heat flux of 3 x 10^8 ev m^{-2} sec^{-1} is obtained for the field tube with foot at 40^o geomagnetic latitude. his flux is very much smaller than the integrated rate of heating per cm^2 olumn between 200 and 1000 km and so has no effect on the thermal structure of he ionosphere.

The foregoing is a review of the steady-state daytime situation. A full analysis of the transient phenomena which occur near sunrise and sunset is beyond the scope of this investigation. However, it was noted in Chapter 5 that the time required for the ionospheric temperature to fall to that of the neutral atmosphere when the production of photoelectrons ceases is of the order of a second. This indicates that very shortly after sunset the ionosphere assumes the role of a heat sink for the thermal energy stored in the protonosphere. One objective of this section is to arrive at an estimate of this thermal energy. Once this has been obtained, a time-dependent expression for the heat flux is derived, and an estimate of the time required for the thermal energy of the protonosphere to drain into this heat sink is obtained. Possible effects on the nighttime ionosphere over this duration of time are then considered.

6.5.1 Thermal Capacity of the Protonosphere

The thermal energy stored in a protonospheric field tube which could be transferred to the neutral atmosphere is given by

$$U = 3Nk\ (T_e - T_n), \tag{6.29}$$

where N is the total electron content of the tube and T_e and T_n are, respectively the common ion-electron temperature and the temperature of the neutral atmosphere.

A figure of 4.52×10^{12} cm^{-2} was quoted in Chapter 2 for the total electron content of the tube with base of unit area on the 1000 km surface at geomagnetic latitude 40^{o} and bounded at the upper end by the equatorial plane. This result was obtained for a distribution of ionization in diffusive equilibrium at a

temperature of $1000^{\circ}K$ with electron density 10^3 cm^{-3} at the equatorial plane. If the temperature is increased to $3000^{\circ}K$ and the electron density at the equatorial plane is held fixed, diffusive equilibrium cannot be restored unless some ionization is removed from the tube. The figure of 4.52×10^{12} cm^{-2} is therefore an overestimate of the total electron content at a temperature of $3000^{\circ}K$. The factor by which it must be multiplied to give the correct value can be found from Equations (2.36) and (2.37) of Chapter 2. Equation (2.36) shows that $N \propto I(T)$ for a fixed density at the equatorial plane; the factor by which N must be multiplied to obtain the total content at temperature T' is therefore $I(T')/I(T)$. From Equation (2.37) it can be seen that the temperature dependence of this ratio is given by Figure 2.4. For temperatures in excess of $1500^{\circ}K$ it is practically independent of temperature and equal to about 0.8. Thus, at a temperature of $3000^{\circ}K$ the total electron content of the field tube is about 3.6×10^{12} cm^{-2}.

The quantity of interest in this section is the thermal capacity U at time of sunset. It will be assumed that $T_e = 2800^{\circ}K$ and that $T_n = 800^{\circ}K$. The latter figure is more appropriate to midnight than time of sunset, and is chosen here only because it is the neutral gas temperature used in the time-dependent calculation to follow. Substitution of these temperatures and the total electron content into Equation 6.29 gives $U \approx 3$ ergs.

This energy is of the order of the thermal energy found in a cm^2 column of the ionosphere during the day. It was pointed out in Chapter 5 that this would be lost to the neutral atmosphere in a matter of seconds in the event that input should cease and that the neutral atmosphere could dispose of this energy in an equivalent time without any alteration of its thermal structure. Hence,

it is clear that at sunset the collapse of ionospheric temperature to that of the neutral atmosphere would be delayed only momentarily by the extra energy contained in the protonosphere. However, the important difference between energy stored in the protonosphere and energy stored in the ionosphere is one of location rather than magnitude. The energy stored in the protonosphere must be transported down to ionospheric heights before it can be lost to the neutral atmosphere. A very short time after sunset the F-region electron temperature is low, while that of the protonosphere is still essentially at its daytime value: a large temperature gradient exists and, because of the large value of the thermal conductivity, there is a considerable flux of energy through the base of the protonosphere. But, as the thermal energy of the protonosphere is drained away, this gradient and the thermal conductivity both decrease. Because of the 5/2 power dependence on the temperature, the thermal conductivity is reduced drastically as the electron temperature falls. The operation of the thermal conductivity here is much like that of a valve, and, after an initial rush of thermal energy, the cooling of a protonospheric field tube can best be described as the result of a "leakage of energy" through its base.

6.5.2 Rate of Cooling of the Protonosphere

In order to obtain an estimate of the time of cooling of a protonospheric field tube, a simplified model will be adopted in which the field tube is replaced by an isothermal heat reservoir of total energy content U. The base of this heat reservoir is at the 1000 km level and thermal contact is established with a heat sink at the 300 km level through the intermediate conducting electron gas. The temperature of the heat reservoir is initially $T_o = 2800^oK$, while

that of the heat sink remains at all times equal to the temperature of the neutral atmosphere, which is 800°K. If the heat reservoir is assumed to have unit area at the base, then energy is lost at a rate

$$\Phi = - \frac{dU}{dt} \quad , \tag{6.30}$$

where Φ is the downward heat flux, given by the steady-state heat conduction equation:

$$\Phi \approx 6.84 \times 10^{5} \; T_e^{5/2} \; \frac{(T_e - T_n)}{\Delta z} \quad . \tag{6.31}$$

The equation for the rate of change of the temperature of the reservoir is obtained from the equality of the right-hand sides of Equations (6.30) and (6.31) after substitution for U from Equation (6.29):

$$3Nk \; \frac{dT_e}{dt} = - 6.84 \times 10^{5} \; T_e^{5/2} \; \frac{(T_e - T_n)}{\Delta z} \quad . \tag{6.32}$$

Upon substitution of $N = 3.6 \times 10^{12}$ cm^{-2} and $\Delta z = 7 \times 10^{7}$ cm and division of both sides by T_n, Equation (6.32) becomes

$$\frac{dx}{dt} = - 9.50 \times 10^{-5} \left\{ x^4 (x^2 - 1) \right\} \quad , \tag{6.33}$$

where $x = (T_e / T_n)^{1/2}$. With $x = x_o$ at time $t = o$, Equation (6.33) can be put into the form

$$\int_{x_o}^{x} \frac{dx}{x^4 (x^2 - 1)} = - 9.50 \times 10^{-5} t$$

The left-hand side of this equation may be integrated by the technique of partial fractions to give

$$\left\{ \frac{1}{3x^3} + \frac{1}{x} - \frac{1}{2} \log \left| \frac{1+x}{1-x} \right| \right\} \Bigg|_{x}^{x_o} = 9.50 \times 10^{-5} \, t \, , \qquad (6.34)$$

an implicit relation for the temperature of the heat reservoir as a function of time.

Values of t for temperatures between 2800°K and 1000°K have been obtained from Equation (6.34), and the result used in conjunction with Equation (6.31) to find the downward flux Φ as a function of time. This function is plotted in Figure 6.6. It can be seen that a large downward flux is present initially. This large flux quickly cools the heat reservoir and lowers the thermal conductivity. According to this model, the flux levels off at a base value near 5×10^7 ev cm^{-2} sec^{-1} in a little over an hour. At this time, the temperature of the heat reservoir is about 1000°K. Substitution of $T_e = 1000^{\circ}K$ into Equation (6.29) shows that the heat reservoir still contains about 1.7×10^{11} ev of energy, and energy may therefore continue to leak into the heat sink at a rate of 5×10^7 ev cm^{-2} sec^{-1} for another hour.

Thus, the idealized model for the process of the nighttime cooling of a protonospheric field tube gives a time constant of the order of two hours for cooling to take place at middle geomagnetic latitudes. This differs considerably from the time constant for cooling of the ionosphere, which is of the order of a second. The difference is due to the fact that the thermal energy stored in the protonosphere cannot be transferred to the neutral atmosphere in situ but must be conducted downward to ionospheric levels at a rate governed

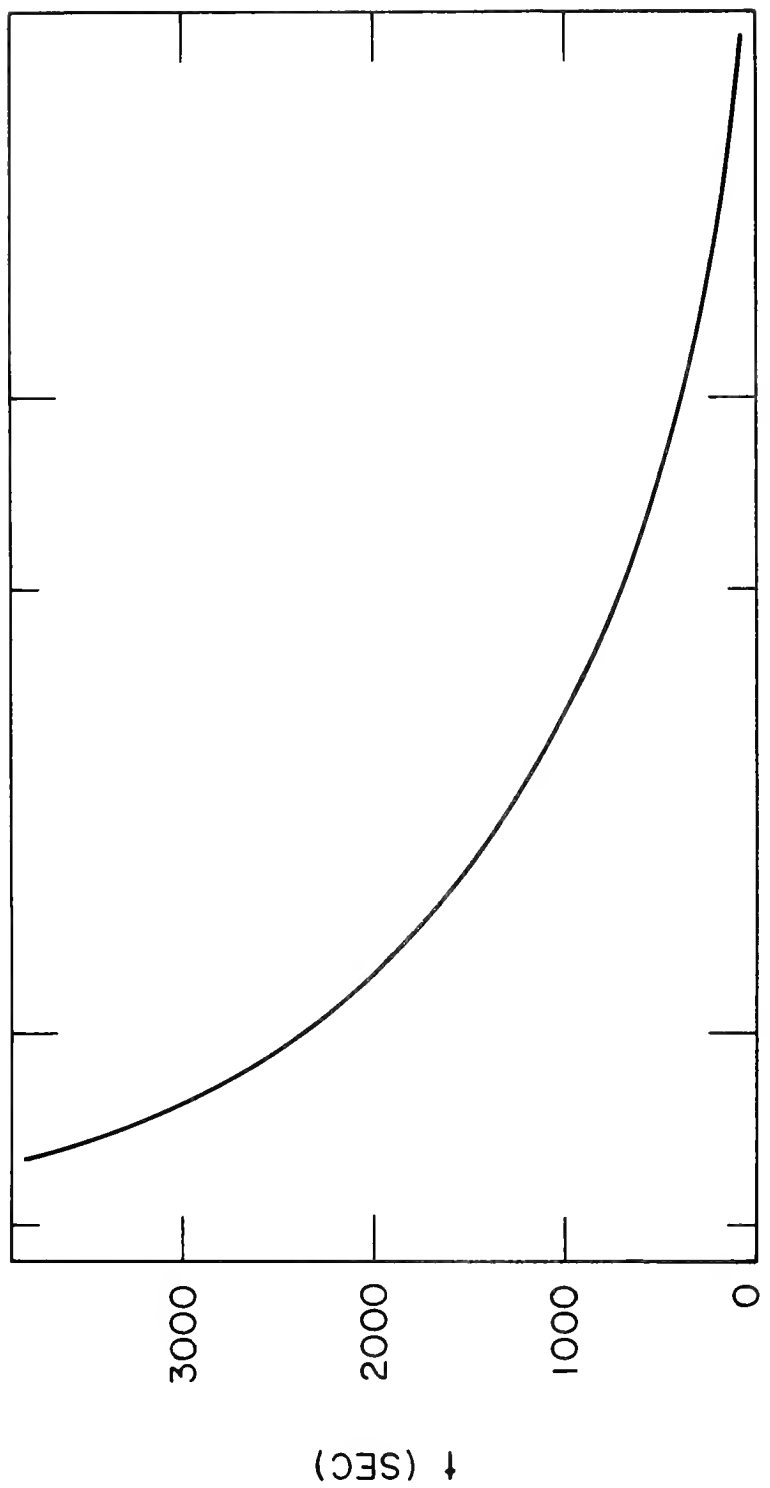

Figure 6.6 Decrease of protonospheric heat flux with time after sunset.

by a thermal conductivity that varies directly as a high power of the temperature.

6.5.3 Effect on the Nighttime F2-Layer

The use of an idealized model for the cooling process leads to an estimate of the time of cooling, but further refinement is necessary before application of the theory to explain observed nighttime phenomena can be attempted. To begin with, replacement of the field tube by a heat reservoir leads to an underestimate of the cooling time. The reason for this is that a heat reservoir responds uniformly and instantaneously to the demand for heat, whereas, in the protonosphere, heat can be conducted through the base of the field tube only as fast as it can be supplied by conduction from higher regions of the tube. It has been shown that during the day this rate is about 3×10^8 ev cm^{-2} sec^{-1}. As this figure will apply also to the time of sunset, the large flux of heat which is initially present for the model used to obtain the cooling time cannot be supplied by the protonosphere, and the time required for the flux to approach the base value has therefore been underestimated.

The duration of the second phase of the cooling process--the time required for the base value flux to conduct away the remainder of thermal energy within the tube--has also been underestimated. This is due to the underestimate of the thermal content of the field tube at this time that results when the temperature at the base of the tube is used. For definiteness, let the time at which the flux levels off at the base value be the time required for the temperature at the base of the protonosphere to fall to $1000^{\circ}K$. At this time, a temperature difference of some $200^{\circ}K$ exists between the 300 and 1000 km levels. In order that the required flux be supplied to the base of the protonosphere,

this gradient must persist well up into the field tube, diminishing gradually with distance above the base. This means that much of the protonosphere is at a temperature considerably in excess of the temperature at the base, and the thermal content at this time is therefore much higher than the estimated value.

Thus, the simplified theory which replaces the field tube by a heat reservoir does not very accurately describe what takes place during the night, and a full time-dependent solution that takes account of the slow response of the field tube is desirable. Such a treatment will not be attempted here. Rather it will be asserted that a downward heat flux of the order of 10^8 ev cm^{-2} sec^{-1} can be maintained throughout most of the night at middle geomagnetic latitudes. In the light of the evidence presented thus far, this is not unreasonable.

The programmed method of solution of the heat conduction equation used to obtain the daytime profiles presented in Chapter 5 has been modified by Bowhill (1964) to handle the nighttime situation. Figure 6.7 is taken directly from this reference and shows the behavior of the electron and ion temperatures for an assumed downward flux of 10^8 ev cm^{-2} sec^{-1} at 1000 km. The model atmosphere is the Harris-Priester model S = 100 for time of local midnight; the temperature of the neutral atmosphere above 300 km is about $720^{\circ}K$. The behavior of the electron temperature profile is similar to that assumed in the calculation of the cooling time: the electron temperature is equal to that of the neutral atmosphere near 300 km and increases with altitude above this level at a constant rate that is such that $T_e \approx 1000^{\circ}K$ at the 1000 km level.

A feature of the temperature profile of Figure 6.7 which deserves special comment is the decrease of the electron temperature to values below that of the neutral atmosphere in the region between 125 and 350 km. This is a consequence

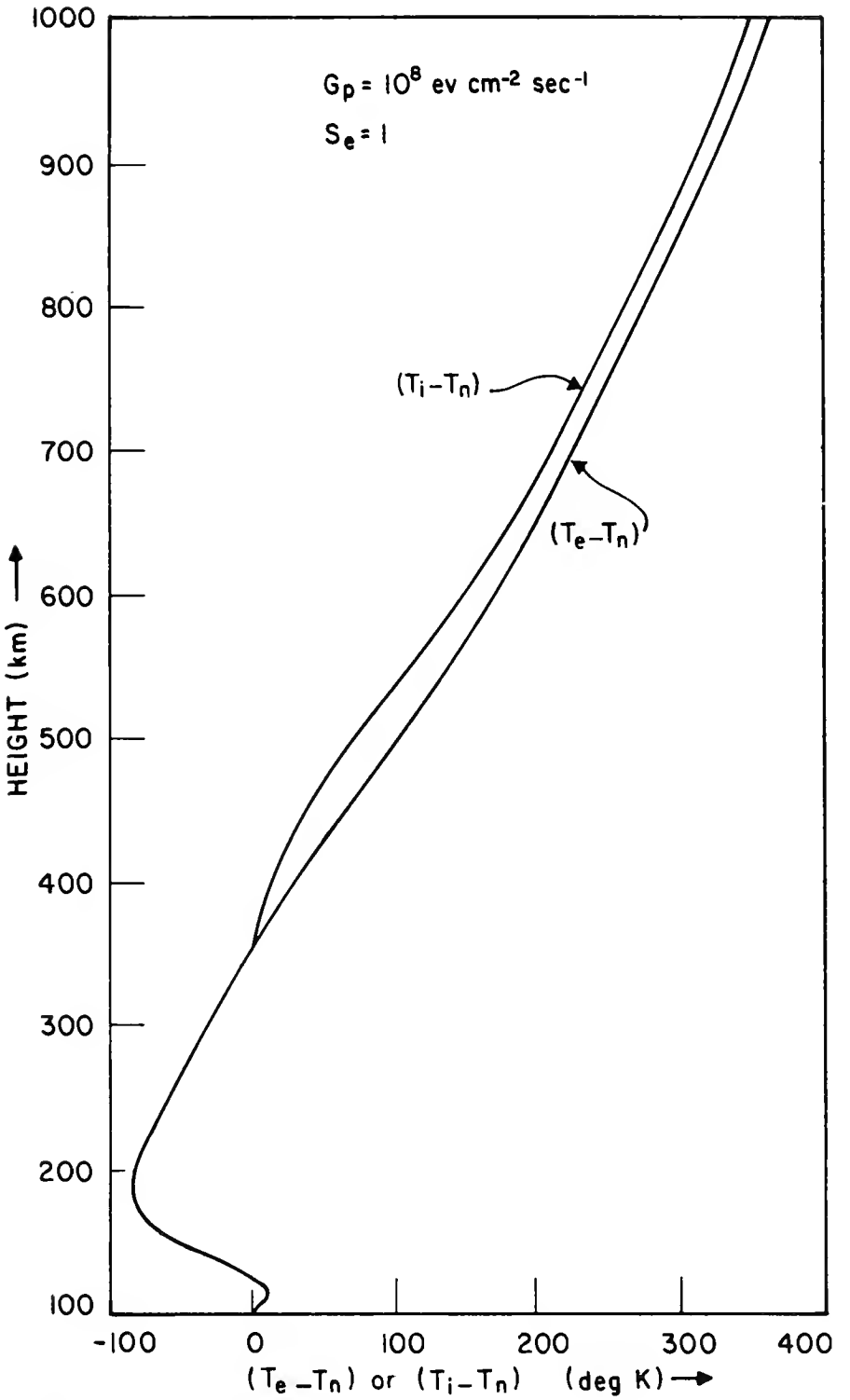

Figure 6.7 Nighttime electron and ion temperatures for the case of a protonospheric heat flux of 10^8 ev cm^{-2}sec^{-1}. (after Bowhill, 1964).

of the steep gradient of neutral gas temperature in the region between 100 and 250 km. The electron temperature stays fairly close to the neutral gas temperature in this altitude range and must also be characterized by a steep gradient. This gradient of electron temperature implies the existence of a downward heat flux through this region, and this can only be maintained if heat is supplied to the electron gas in the upper part of the region and given up by the electron gas in the lower part of the region. Because of the rapid decrease of neutral concentration with altitude (and consequent decrease of efficiency of heat transfer), the region in which energy is gained by the electron gas will be the more extensive of the two and will be characterized by a larger temperature difference. Inspection of Figure 6.7 shows that the electron gas gains energy throughout the region between 125 and 350 km, conducting it downward through the 125 km level, with subsequent transfer to the neutral atmosphere between 125 and 100 km. The downward heat flux reaches a minimum value of 5.6×10^{7} ev cm^{-2} sec^{-1} at 355 km. Below this level the heat flux increases with decreasing altitude, reaching a maximum value of 1.5×10^{8} ev cm^{-2} sec^{-1} at the 125 km level. It should be noted that at all altitudes below about 350 km the ion temperature is equal to that of the neutral atmosphere. At higher altitudes, the ion gas loses thermal contact with the neutrals, and the ion temperature is close to that of the electron gas above 600 km.

The gradient of electron temperature above 350 km is very close to that which has been deduced from the Ariel I satellite measurements (Willmore *et al.*, 1964) of electron temperature near local midnight. It may therefore be concluded that the heat flux implied by the variation with altitude of the night-

time electron temperature measured by Ariel I is supplied by the protonosphere and that it is not necessary to invoke some other source of heating, such as a flux of energetic particles.

The interesting suggestion has been put forward by Bowhill (1964) that relatively small horizontal inhomogeneties in the magnitude of the nighttime protonospheric heat flux can account for the existence of weak field-aligned irregularities that are observed in the neighborhood of the F2-layer maximum. Such an effect is the result of the sensitivity of the electron temperature at and above the F2-layer maximum to changes in the heat flux coupled with the response of electron density to changes in electron temperature.

The argument of Bowhill (1964) will be recapitulated briefly here. Consider first the response of electron density to temperature change. Most of the ionization above the F2-layer maximum is in diffusive equilibrium and the plasma pressure p_p at height z_o may therefore be written as the sum of the partial pressures of the electron and ion gases:

$$p_p(z_o) = k(T_e + T_i)/n_e(z_o) \quad . \tag{6.35}$$

This may be equated to the total weight of the plasma above z_o:

$$p_p(z_o) = m_i \bar{g} \int_{z_o}^{\infty} n_e(z)\, dz \quad , \tag{6.36}$$

where m_i is the mass of an oxygen ion and $\bar{g}$ is the mean gravitational acceleration in the region considered. It is then argued that a change of electron

temperature is not likely to affect the total content above z_o to any great extent, since the rate at which ionization disappears is primarily a function of ion and neutral gas temperatures. The right-hand side of Equation (6.36) therefore remains constant with change of electron temperature, and it follows from Equation (6.35) that

$$n_e (z_o) \propto (T_e + T_i)^{-1} . \qquad (6.37)$$

The sensitivity of electron temperature to changes in the protonospheric heat flux has been determined from integrations of the heat conduction equation using the modified nighttime program. When this result is combined with the temperature dependence expressed by Equation (6.37), the response of the electron density to changes in the flux can be found. The results of this analysis show that a local change in protonospheric heat flux from 10^8 to 3×10^8 ev cm^{-2} sec^{-1} would be sufficient to produce a field-aligned decrease of nearly 10 per cent in the peak electron density of the F2-layer. At much higher altitudes (say 1000 km) this effect is reversed, however, and a field-aligned increase is produced. For the change of flux just quoted this increase may exceed 30 per cent, depending on the concentration of helium ions present.

7. SUMMARY AND CONCLUSIONS

It was proposed in the introductory chapter that a study be made of the effect of diffusion of ionization across the 1000 km surface on the morphology of the F2-layer. Production and loss of ionization are negligible above 1000 km and the region above this level may be regarded as a reservoir of ionization. As the propagation of atmospheric whistlers is strongly influenced by the quantity and the state of the ionization in this region, the investigation was started with a calculation of the relation between the whistler dispersion, the temperature of the protonosphere and the ionization content of a protonospheric field tube.

A steady-state model of the F2-layer was set up in order to estimate the magnitude of the flow of ionization into or out of the topside of the layer that is required to produce a significant effect on the structure of the layer. This was followed by a study of the processes which account for this flow of ionization: diffusion of hydrogen ions through the topside and charge exchange between hydrogen ions and atomic oxygen. The results of this analysis were then used to set up a steady-state model for ionosphere-protonosphere coupling. It was found that the rate of exchange of ionization is strongly dependent upon the daytime ion temperature; this led to an investigation of the thermal structure of the ionosphere.

The processes which heat and cool the ionosphere were discussed in full detail. An expression was derived for the rate of heat input in the topside ionosphere which includes the contribution from photoelectrons which originate at and below the F2-layer maximum. Daytime electron and ion temperature

profiles were obtained over the altitude range between 120 and 1000 km from solutions of the steady-state heat conduction equation. A detailed comparison with observation was presented.

The theory developed for the rate of heat input to the topside ionosphere was extended to derive the rate of heat input to the ionization contained within a protonospheric field tube by photoelectrons which originate at F2-layer heights. Under the assumption that the ionization within the field tube can be cooled only by conduction of heat downward through the 1000 km surface, the temperature profile along the length of the field tube was calculated. The effect of this downward flux of heat on the daytime ionosphere was described.

The protonospheric temperature thus derived was used to estimate the thermal energy stored in a protonospheric field tube. Nighttime electron and ion temperature profiles were derived under the assumption that thermal energy stored in the protonosphere is the only source of heat input to the nighttime ionosphere. Effects on the nighttime ionosphere which may result from small horizontal inhomogeneities in the rate at which heat is thus supplied were described.

Thus, in summary, an investigation which began as an inquiry into the effect on the F2-layer of diffusion of ionization across the 1000 km surface has led ultimately to a study of the thermal structure of both the ionosphere and the protonosphere. Conclusions based on this investigation are summarized in the succeeding sections.

7.1 Interpretation of the Whistler Dispersion

The analysis of the protonosphere has been carried out with reference to the ionization contained within a tube of magnetic force that intersects unit area on the 1000 km surface at latitude θ and is bounded at the other end by the

plane of the geomagnetic equator. The ionization within this protonospheric field tube is distributed hydrostatically along the magnetic field line that defines the axis of the tube. The electron density at the equatorial face of the tube may change either as a result of a redistribution of ionization along this field line following a temperature change or as a result of ionization entering or leaving the tube.

The electromagnetic disturbance which produces an atmospheric whistler propagates along the line of force of the geomagnetic field. The integrand of the equation which defines the whistler dispersion is a weighted average of the electron density along the path of propagation of the disturbance, with the largest contribution coming from the region near the plane of the geomagnetic equator. It is for this reason that the whistler dispersion is commonly regarded as a measure of electron density in the equatorial plane. However, the analysis of Chapter 2 shows that the integrand is sufficiently sensitive to changes in electron density at other points along the field tube that changes in the electron density at the equatorial face of the tube due to temperature changes alone do not affect the value of the integral. It is concluded from this that the whistler dispersion is a measure of total electron content of the field tube rather than the electron density at the equatorial face of the tube.

The main difficulty in the interpretation of whistler data is identification of the field line along which the disturbance producing the whistler propagates. Considerable progress has been made in the past few years in this area and in improvement of the techniques of recording whistlers. It may be expected that reliable measurements of both diurnal and seasonal variation of

field tube content will be forthcoming from whistler observations. Present inferences of variation of field tube content are largely qualitative, however; as indicated in Chapter 2, the most that can be said presently is that only a rather small fraction of the ionization contained in a middle-latitude field tube participates in the diurnal exhange process with the ionospheric F2-layer.

7.2 Exchange of Ionization Between the Ionosphere and the Protonosphere

A steady-model of the F2-layer was obtained in Chapter 3 by solution of the steady-state continuity equation in an assumed model atmosphere. The solution was obtained subject to the upper boundary condition that the flux tends to a finite limiting value at infinite heights and therefore represents the distribution of ionization in an F2-layer characterized by an influx or efflux of ionization at the top of the layer. This solution was used to obtain a numerical estimate of the flux of ionization at the top of the layer that is needed to account for two well-known features of the layer: the seasonal anomaly and the persistence of the layer at night.

Diffusion of atomic hydrogen ions from the base of the protonosphere downward through the topside, attended by conversion of these ions to atomic oxygen ions through the chemical process of charge exchange with atomic oxygen, results in a downward flux of atomic oxygen ions into the F2-layer. Because the ionization potentials of atomic hydrogen and atomic oxygen are nearly the same, the charge exchange reaction proceeds almost equally rapidly in either direction. This allows the chain of events just described to proceed in reverse order, a flow of atomic hydrogen ions into the protonosphere being maintained at the expense of the atomic oxygen ions of the F2-layer.

A diurnal exchange of ionization between the F2-layer and the protonosphere results from the large diurnal variation of F2-layer content. This variation is such that ionization flows into the layer at night.

The exchange process was analyzed in Chapter 4 by solution of the steady-state form of the continuity equation that governs the distribution of ionized hydrogen in the region of the topside where these ions are a minor constituent. A feature of this solution is the existence of limit to the rate at which ionized hydrogen may diffuse upward through atomic oxygen ions and into the protonosphere. This limit, called the critical flux, is independent of any assumption of protonospheric field tube content. On the basis that this quantity is some two orders of magnitude smaller than the flux needed to account for the seasonal anomaly, it is concluded that the process of diffusive transport of ionization along magnetic field lines from the summer hemisphere to the winter hemisphere cannot be responsible for this feature of F2-layer behavior.

A flux of ionization downward into the F2-layer at night cannot long be maintained at a value in excess of the daytime critical flux, since the ionization thus removed from a protonospheric field tube must be replenished during the day. In order to establish whether or not the rate of exchange of ionization is close to this upper limit, a steady-state model for ionosphere-protonosphere coupling has been set up using the observed day to night ratio of F2-layer density and the day to night ratio of the concentration of ionized hydrogen at the base of the protonosphere calculated in Chapter 2. The results obtained from this model indicate that the average rate of exchange of ionization is very close to the upper limit set by the critical flux. The concentration at the base of the

protonosphere appears as an eigenvalue in the calculations; the value deduced from the assumed parameters of the model is in good agreement with the value inferred from middle-latitude whistler data.

The results obtained from the steady-state model of the coupling process form the basis for the conclusion that the magnitude of the downward flux into the nighttime F2-layer is equal to that of the critical flux. In order to determine whether or not a downward flux is responsible for maintenance of the nighttime F2-layer it is necessary to compare the figure for the required flux, deduced in Chapter 3, with the value of the critical flux. Comparison with the value of the critical flux calculated in Chapter 4 shows that this quantity is a factor of 20 smaller than the flux required to maintain the layer. Subsequent analysis of the thermal structure of the ionosphere has shown, however, that the temperature used in the calculation of the critical flux is too small by a factor of 2. As the temperature enters the expression for the critical flux raised to the 3/2 power, the discrepancy is reduced to a factor of 7.

On the basis that the current estimates of the magnitudes of physical quantities which appear in the expression for the critical flux give a value that is less by a factor of 7 than the required flux, the conclusion is drawn that the downward flux of ionization into the nighttime F2-layer cannot be solely responsible for the maintenance of the layer. This conclusion should, however, be accepted with some reservation until further research has confirmed that the value of the diffusion coefficient and the value of the concentration of atomic hydrogen adopted for the calculation of the critical flux are correct.

7.3 The Thermal Structure of the Ionosphere

The investigation of the thermal structure of the ionosphere began with an analysis of ionospheric heating. Particular attention was given to the

phenomenon of non-local heating. It was first shown that this results in an excess of the rate of heat input over that deduced from the assumption of local heating at altitudes above the photoelectron escape level. This excess increases with altitude, with the result that the non-local heating rate exceeds that of the local heating rate by an order of magnitude near the base of the protonosphere. Subsequent analysis showed, however, that electron temperature profiles calculated with the local and with the non-local heat input functions are practically indistinguishable, a result which is attributed to the effects of thermal conduction.

Previous theoretical investigations of the lack of thermal equilibrium within the ionosphere have all mentioned the possible influence of thermal conduction, but the calculations of electron temperature profiles have been carried out without taking explicit account of this phenomenon. This, coupled with the growing number of experimentally obtained temperature profiles that are at variance with those predicted by the theory, has been the motivation for a treatment that includes heat transport by thermal conduction. The results of this analysis indicate that effects of thermal conduction are not negligible particularly near the minimum of the solar cycle, and that the discrepancy between theory and observation is removed when allowance for these effects is incorporated into the theory.

In particular the present analysis shows that.

(1) Electron temperature profiles near sunspot minimum should show a monotonic increase with altitude above 120 km. This is a direct consequence of the relatively low electron content of the ionosphere characteristic of this time. a runaway condition exists and the temperature

profile is set by the requirements of thermal conduction.

(2) Near sunspot minimum the ion temperature is sensibly equal to that of the neutral atmosphere below 500 km; above 900 km it is sensibly equal to the electron temperature.

(3) The electron temperature in the F-region and above is very sensitive to changes in the incoming solar flux near sunspot minimum.

(4) At sunspot maximum the electron temperature reaches a maximum near 225 km, and near-thermal equilibrium exists at and above 350 km. This is a direct consequence of the relatively high electron content characteristic of this time: a runaway condition is encountered only below 250 km and this is kept in check by the large concentration of neutrals found at these low altitudes.

(5) The difference between electron temperature and neutral gas temperature is just such that the daytime electron temperature shows little or no variation over the solar cycle, remaining at a value near 2500^{o}K. Ion temperatures in the topside ionosphere are generally higher at sunspot maximum than at sunspot minimum but the difference is not large. This relative constancy of the topside plasma temperature suggests that changes in the structure of the topside ionosphere observed over the solar cycle are due to changes in chemical composition.

7.4 The Thermal Structure of the Protonosphere

The upward flux of photoelectrons which accounts for the phenomenon of non-local heating in the topside ionosphere is probably the only significant source of protonospheric heating. Calculations based on the production rate used in the calculation of ionospheric electron temperatures and an assumed initial

energy of 10 ev per photoelectron result in a figure of 3×10^8 ev cm^{-2} sec^{-1} for the rate of heat input to the protonospheric field tube with foot at latitude 40^o and bounded by the equatorial plane. The contribution of photoelectrons which escape from the conjugate ionosphere has been included in the calculation. This figure applies to sunspot minimum conditions; it is estimated to be larger by a factor of 5 at sunspot maximum.

A calculation of the temperature distribution within the field tube was carried out under the assumption that cooling of the field tube is effected only by downward conduction of heat. The temperature at the base of the field tube was taken to be the 1000 km temperature obtained from the ionospheric electron temperature calculations; figures of 2800^oK and 2500^oK apply, respectively, to the minimum and maximum of the solar cycle. The calculated temperature distributions are such that temperatures of 3000^oK and 3400^oK are found at the equatorial face of the tube with foot at latitude 40^o for the respective extremes of the solar cycle.

From this the conclusion may be drawn that, for low and medium sunspot numbers, heating by the photoelectrons which escape the F2-layer does not significantly increase the temperature of the protonosphere over that of the topside ionosphere for latitudes of 40^o and less. This is a consequence of the high thermal conductivity of an overheated electron gas: the principal cause of the high electron and ion temperatures found in the protonosphere is the high electron temperature at F2-layer heights that is necessary to maintain the steep electron temperature gradient in the lower F-region.

7.5 Exchange of Energy Between the F2-layer and the Protonosphere

It has been estimated that escape of photoelectrons from unit column of the ionosphere supplies an upward flux of 10^8 cm^{-2} sec^{-1} across the 1000 km

surface at sunspot minimum. The mean energy of these photoelectrons is a fraction 8/15 of the initial energy, or about 6 ev for the assumed initial energy of 10 ev. This flux of photoelectrons supplies heat to both halves of the protonospheric field tube with foot at latitude 40^{o} at a total rate of 3×10^{8} ev cm^{-2} sec^{-1}. The remaining 3×10^{8} ev cm^{-2} sec^{-1} is carried downward in the form of kinetic energy across the 1000 km surface at the conjugate end of the field tube and ultimately appears as thermal energy in the conjugate ionosphere.

It has been shown in Chapter 5 that a heat input rate of this order of magnitude has no effect on the electron temperature in the daytime ionosphere. It should be noted, however, that the time of sunrise and the time of sunset may each differ considerably at magnetic conjugate points. A heat input of 3×10^{8} ev cm^{-2} sec^{-1} may be expected to have a significant effect on the temperature and the structure of the ionosphere at the dark end of the field tube. This is particularly so near sunspot maximum when the above figure must be increased by a factor of 5.

For the case of complete symmetry with respect to the geomagnetic equator, the escape fluxes from both hemispheres combine to supply heat to the field tube with foot at latitude 40^{o} and bounded by the equatorial plane at a rate of 3×10^{8} ev cm^{-2} sec^{-1} at sunspot minimum. This results in a downward flux of thermal energy of the same magnitude through the 1000 km surface. This heat flux has no effect on the thermal structure of the sunspot minimum ionosphere during the day. The same conclusion has been reached for the sunspot maximum case.

This conclusion does not apply to the nighttime ionosphere. An estimate of the thermal energy content of the protonospheric field tube has been

combined with a time-dependent model of the nighttime cooling process to show that a heat flux of about 10^8 ev cm^{-2} sec^{-1} downward through the 1000 km surface may be expected to persist throughout the night. Comparison of the electron temperature profile through the topside ionosphere deduced from this heat input with nighttime temperature measurements supports the conclusion that this is the only significant source of heating of the nighttime ionosphere. It has also been found that the temperature near and above the F2-layer maximum is sufficiently sensitive to the magnitude of the heat flux that small horizontal inhomogeneities in the flux result in field-aligned irregularities of electron density that closely resemble those which are often observed near the maximum of the nighttime layer.

7.6 Suggestions for Further Research

The effect that a departure from thermal equilibrium has on the structure of the ionosphere should be investigated. In particular, a study should be made of: (1) the effect of the high electron temperature on the chemical processes in the lower F-region and in the E-region; (2) the structure of the ionosphere near sunrise and sunset, when large changes in electron and ion temperature occur, (3) the effect of the large ion temperature gradient on the structure of the topside.

The conclusions concerning the thermal structure of the protonsophere and the thermal coupling between the ionosphere and protonosphere are based on calculations for a field tube with foot at 40^o geomagnetic latitude. It would be desirable to extend this analysis to higher latitudes.

The conclusion that the protonosphere is the source of nighttime ionospheric heating is based on a simple model of the cooling of a protonospheric field tube. A more realistic description of this process is needed.

REFERENCES

Allcock, G. McK., "The Electron Density Distribution in the Outer Ionosphere Derived from Whistler Data," J. Atmosph. Terr. Phys., 14, 185-199, 1959.

Bates, D. R., and Patterson, T. N. L., "Hydrogen Atoms and Ions in the Thermosphere and Exosphere," Planet. Space Sci., 5, 257-273, 1961.

Bates, D. R., and Patterson, T. N. L., "Helium Ions in the Upper Atmosphere," Planet. Space Sci., 9, 599-605, 1962.

Bourdeau, R. E., Whipple, E. C., Donley, J. L., and Bauer, S. J., "Experimental Evidence for the Presence of Helium Ions Based on Explorer VIII Satellite Data," J. Geophys. Res., 67, 467-476, 1962.

Bowen, P. J., Boyd, R. L. F., Raitt, W. J., and Willmore, A. P., "Ion Composition in the Upper F-region," Proc. Roy. Soc. A, 1964, (in the press).

Bowhill, S. A., "The Formation of the Daytime Peak of the Ionospheric F2-Layer," J. Atmosph. Terr. Phys., 24, 503-519, 1962a.

Bowhill, S. A., "An Accretive Model for the Ionospheric F2-Layer," Paper presented at USNC-URSI meeting, Ottawa, Canada, October 15-17, 1962b.

Bowhill, S. A., "Origin of Field-Aligned Irregularities in the F2-Layer," Paper presented at AGARD Symposium on Spread F, Copenhagen, Denmark, August 26-29, 1964, (to be published).

Bowles, K. L., Ochs, G. R., and Green, J. L., "On the Absolute Intensity of Incoherent Scatter Echoes from the Ionosphere," J. Res. NBS, 66D, 395-407, 1962.

Brace, L. H., Spencer, N. W., and Carignan, G. R., "Ionosphere Electron Temperature Measurements and Their Implications," J. Geophys. Res., 68, 5397-5412, 1963.

Brown, E. C., "Basic Data of Plasma Physics," Technology Press and Wiley, 1959.

Butler, S. T., and Buckingham, M. J., "Energy Loss of a Fast Ion in a Plasma," Phys. Rev., 126, 1-4, 1962.

Carpenter, D. L., "Electron Density Variations in the Magnetosphere Deduced from Whistler Data," J. Geophys. Res., 67, 3345-3360, 1962.

Carpenter, D. L., and Smith, R. L., "Whistler Measurements of Electron Density in the Magnetosphere," Revs. Geophys., 2, 415-439, 1964.

Chandra, S., "Electron Density Distribution in the Upper F-Region," J. Geophys. Res., 68, 1937-1942, 1963.

Chapman, S., "The Absorption and Dissociative or Ionizing Effect of Monochromatic Radiation in an Atmosphere on a Rotating Earth," Proc. Phys. Soc., 43, 484-501, 1931.

Chapman, S., and Cowling, T. G., "The Mathematical Theory of Non-Uniform Gases," Cambridge University Press, 1958.

Croom, S. A., Robbins, A. R., and Thomas, J. O., "Tables of Ionosphere Electron Density," Cavendish Laboratory, Cambridge, 1959.

Croom, S. A., Robbins, A. R., and Thomas, J. O., "Variation of Electron Density in the Ionosphere with Magnetic Dip," Nature, 185, 902-903, 1960.

Dalgarno, A., "Ambipolar Diffusion in the F-Region," J. Atmosph. Terr. Phys., 12, 219, 1958.

Dalgarno, A., "Ambipolar Diffusion in the F-Region," J. Atmosph. Terr. Phys., 26, 939, 1964.

Dalgarno, A., McElroy, M. B., and Moffett, R. J., "Electron Temperatures in the Ionosphere," Planet. Space Sci., 11, 463-484, 1963.

Dalgarno, A., and Moffett, R. J., "Electron Cooling in the D-Region," Planet. Space Sci., 9, 439-441, 1962.

Damon, K. R., Hall, L. A., and Hinteregger, H. E., "Solar Extreme Ultraviolet Photon Flux Measurements in the Upper Atmosphere of August, 1961," Space Research III, edited by W. Priester, 745-759, North-Holland Publishing Co., Amsterdam, 1963.

Donley, J. L., "Experimental Evidence for a Low Ion-Transition Altitude in the Upper Nighttime Ionosphere," J. Geophys. Res., 68, 2058-2060, 1963.

Dungey, J. W., "Electrodynamics of the Outer Atmosphere," The Physics of the Ionosphere, 229-236, The Physical Society, London, 1955.

Dwight, H. B., "Tables of Integrals and Other Mathematical Data," Macmillan Co., 1961.

Evans, J. V., "Diurnal Variation of the Temperature of the F-Region," J. Geophys. Res., 67, 4914-4920, 1962.

Evans, J. V., "Ionospheric Temperatures during the Launch of NASA Rocket 8.14 on July 2, 1963," J. Geophys. Res., 69, 1436-1444, 1964.

Evans, J. V., and Loewenthal, M., "Ionospheric Backscatter Observations," Planet. Space Sci., 12, 915-944, 1964.

Ferraro, V. C. A., private communication, 1964.

Gliddon, J. E. C., and Kendall, P. C., "The Effects of Diffusion and of Attachmentlike Recombination on the F2-Region," J. Geophys. Res., 65, 2279-2284, 1960.

Gordon, W. E., "Incoherent Scattering of Radio Waves by Free Electrons with Application to Space Exploration by Radar," Proc. IRE, 46, 1824-1829, 1958.

Hale, L. C., "Ionospheric Measurements with a Multigrid Retarding Potential Analyzer," J. Geophys. Res., 66, 1554, 1961.

Hanson, W. B., "Upper-Atmosphere Helium Ions," J. Geophys. Res., 67, 183-188, 1962.

Hanson, W. B., "Electron Temperatures in the Upper Atmosphere," Space Research III, 282-302, edited by W. Priester, North-Holland Publishing Company, Amsterdam, 1963.

Hanson, W. B., and Johnson, F. S., "Electron Temperatures in the Ionosphere," Memoires Soc. R. Liege, IV, 390-423, 1961.

Hanson, W. B., and Ortenburger, I. B., "The Coupling between the Protonosphere and the Normal F-Region," J. Geophys. Res., 66, 1425-1435, 1961.

Hanson, W. B., and Patterson, T. N. L., "Diurnal Variation of the Hydrogen Concentration in the Exosphere," Planet. Space Sci., 11, 1035-1052, 1963.

Hanson, W. B., Patterson, T. N. L., and Degaonkar, S. S., "Some Deductions from a Measurement of the Hydrogen Ion Distribution in the High Atmosphere," J. Geophys. Res., 68, 6203-6205, 1963.

Harris, I., and Priester, W., "Theoretical Models for the Solar-Cycle Variation of the Upper Atmosphere," NASA Technical Note D-1444, 1962.

Harris, I., and Priester, W., "Relation between Theoretical and Observational Models of the Upper Atmosphere," J. Geophys. Res., 68, 5891-5894, 1963.

Hines, C. O., "The Upper Atmosphere in Motion," Quart. J. R. Met. Soc., 89, 1-42, 1963.

Hunt, D. C., "Thermal Balance in the F-Region of the Atmosphere," NBS Technical Note 162, 1962.

Johnson, F. S., "The Ion Distribution above the F2 Maximum," J. Geophys. Res., 65, 577-584, 1960.

Johnson, F. S., "Physics of the Distribution of Ionized Particles in the Exosphere," Electron Density Profiles in the Ionosphere and Exosphere, 404-413, edited by B. Maehlum, Macmillan Co., 1962.

King, J. W., Smith, P. A., Eccles, D., and Helm, H., "The Structure of the Upper Ionosphere as Observed by the Topside Sounder Alouette," D.S.I.R. Doc. I. M. 94, Slough, 1963a.

King, J. W., Eccles, D., Smith, P. A., Dannahy, P., Legg, A., Olatunji, E. O., Rice, K., Webb, G., and Williams, M., "Further Studies of the Topside Ionosphere Based on the Topside Sounder Data," D.S.I.R. Doc. I. M. 112, Slough, 1963b.

Kockarts, G., and Nicolet, M., "Le Probleme Aéronomique de l'Helium et de l'Hydrogen Neutres," Ann. Geophys., 18, 269-290, 1962.

Mange, P., "The Distribution of Minor Ions in Electrostatic Equilibrium in the High Atmosphere," J. Geophys. Res., 65, 3833-3834, 1960.

Mange, P., "Diffusion in the Thermosphere," Ann. Geophys., 17, 277-290, 1961.

Mariani, F., "Distribution of the Photoelectrons and Origin of the Geomagnetic Anomaly in the F2-Layer," J. Geophys. Res., 69, 556-559, 1964.

Nagy, A. F., Brace, L. H., Carignan, G. R., and Kanal, M., "Direct Measurements Bearing on the Extent of Thermal Nonequilibrium in the Ionosphere," J. Geophys. Res., 68, 6401-6412, 1963.

Nicolet, M., "The Aeronomic Problem of Helium," Ann. Geophys., 13, 1-21, 1957.

Nicolet, M., "Helium, an Important Constituent in the Lower Exosphere," J. Geophys. Res., 66, 2263-2264, 1961.

Nisbet, J. S., and Quinn, T. P., "The Recombination Coefficient of the Nighttime F-Layer," J. Geophys. Res., 68, 1031-1038, 1963.

Outsu, J., and Iwai, A., "Ionization in the Outer Atmosphere Inferred from Whistling Atmospherics," J. Geomag. and Geoelec., 10, 135-142, 1959.

Rapp, D., "Accidentally Resonant Asymmetric Charge Exchange in the Protonosphere," J. Geophys. Res., 68, 1773-1775, 1963.

Rishbeth, H., and Barron, D. W., "Equilibrium Electron Distributions in the Ionospheric F2-Layer," J. Atmosph. Terr. Phys., 18, 234-252, 1960.

Rishbeth, H., "Atmospheric Composition and the F-Layer of the Ionosphere," Planet. Space Sci., 9, 149-152, 1961.

Rishbeth, H., "Further Analogue Studies of the Ionospheric F-Layer," Proc. Phys. Soc. London, 81, 65-77, 1963.

Rishbeth, H., "A Time-Varying Model of the Ionospheric F2-Layer," J. Atmosph. Terr. Phys., 26, 657-685, 1964.

Rivault, R., and Corcuff, Y., "Recherche du Point Conjugué Magnétique de Poitiers -- Variation Nocturne de la Dispersion des Sifflements," Ann. Geophys., 16, 530-544, 1960.

Rothwell, P., "Diffusion of Ions between F-Layers at Magnetic Conjugate Points," Proceedings of the International Conference on the Ionosphere, 217-221, Institute of Physics and the Physical Society, London, 1962.

Smith, E. L., "Properties of the Outer Ionosphere Deduced from Nose Whistlers," J. Geophys. Res., 66, 3709-3716, 1961.

Spencer, N. W., Brace, L. H., and Carignan, G. R., "Electron Temperature Evidence for Nonthermal Equilibrium in the Ionosphere," J. Geophys. Res., 67, 157-175, 1962.

Spitzer, L., "Physics of Fully Ionized Gases," 144, Interscience Publishers, 1962.

Stebbings, R. F., Smith, A. C. H., and Ehrhardt, H., "Charge Transfer between Oxygen Atoms and O^+ and H^+ Ions," J. Geophys. Res., 69, 2349-2356, 1964.

Storey, L. R. O., "An Investigation of Whistling Atmospherics," Phil. Trans. Roy. Soc. London, 246, 113-141, 1953.

Taylor, H. A., Brace, L. H., Brinton, H. C., and Smith, C. R., "Direct Measurements of Helium and Hydrogen Ion Concentration and Total Ion Density to an Altitude of 940 Kilometers," J. Geophys. Res., 68, 5339-5348, 1963.

Thomas, J. O., Haselgrove, J., and Robbins, A. R., "Tables of Ionospheric Electron Density," Cavendish Laboratory, Cambridge, 1957.

Thomas, J. O., and Sader, A. Y., "Electron Density at the Alouette Orbit," J. Geophys. Res., 69, 4561-4582, 1964.

Watanabe, K., and Hinteregger, H. E., "Photoionization Rates in the E and F Regions," J. Geophys. Res., 67, 999-1006, 1962.

Willmore, A. P., "Ionospheric Heating in the F-Region," Proc. Roy. Soc. A, 281, 140-149, 1964.

Willmore, A. P., Henderson, G. L., Boyd, R. L. F., and Bowen, P. J., "Electron Temperature in the Upper F-Region," Proc. Roy. Soc. A, 1964, (in the press).

Wright, J. W., "A Model of the F-Region above $h_m F2$," J. Geophys. Res., 65, 185-191, 1960.

Wright, J. W., "Diurnal and Seasonal Changes in Structure of the Mid-Latitude Quiet Ionosphere," J. Res. NBS, 66D, 297-312, 1962.

Yonezawa, T., "On the F-Region Theory," J. Radio Res. Lab., 3, 1-16, 1956.

Yonezawa, T., "On the Influence of Electron-Ion Diffusion Exerted Upon the Formation of the F2-Layer," J. Radio Res. Lab., 5, 165-187, 1958.

Yonezawa, T., "On the Seasonal and Non-Seasonal Annual Variations and the Semi-Annual Variation in the Noon and Midnight Electron Densities of the F2-Layer in Middle Latitudes," J. Radio Res. Lab., 6, 293-309, 1959a.

Yonezawa, T., "On the Seasonal and Non-Seasonal Annual Variations and the Semi-Annual Variation in the Noon and Midnight Electron Densities of the F2-Layer in Middle Latitudes II," J. Radio Res. Lab., 6, 651-668, 1959b.